

RIVISTA INTERNAZIONALE
DI
SCIENZE ECONOMICHE
E
COMMERCIALI

Anno XXXIII

Dicembre 1986

N. 12

Pubblicazione mensile - Sped. in abb. postale, gruppo 111/70 Bologna

SOMMARIO

- Thorstein Veblen: The Theory of Evolutionary Economics as a Social Science (Thorstein Veblen: la teoria evoluzionistica dell'economia come scienza sociale) ROBERT GRIFFIN Pag. 1145
- Proprietà dell'equilibrio di Nash in modelli di competizione tecnologica (Properties of Nash Equilibrium in R&D Models) FLAVIO DELBONO » 1167
- Causes and Consequences of Labor Militancy (Cause e conseguenze della militanza sindacale) MANFRED GÄRTNER and ERWIN HERI » 1185
- Direct and Indirect Labor and Demand for Workers and Working Hours (Lavoro diretto e indiretto in un modello di domanda di lavoro) PAOLO PALAZZI and PAOLO PIACENTINI » 1207
- La dimensione aziendale nella teoria dell'impresa (Size in the Theory of the Firm) ROBERTO BUSON » 1219
- I grandi saggi di Giuseppe Di Nardi (Di Nardi's Essays) ACHILLE AGNATI » 1233

SOTTO GLI AUSPICI DELLA
UNIVERSITÀ COMMERCIALE LUIGI BOCCONI
E DELLA UNIVERSITÀ DEGLI STUDI DI MILANO

CEDAM - CASA EDITRICE DOTT. A. MILANI - PADOVA

COMITATO DI DIREZIONE - EDITORIAL BOARD

HENRI BARTOLI (Université de Paris) - WILLIAM J. BAUMOL (Princeton University) - FEDERICO CAFFÈ (Università di Roma) - GIOVANNI DEMARIA (Accademia Nazionale dei Lincei) - WILLIAM D. GRAMPP (Illinois University) - ARNALDO MAURI (Università di Milano) - ARIBERTO MIGNOLI (Università Bocconi) - ANTONIO MONTANER (Universität Mainz) - HISAO ONOE (Kyoto University) - ALBERTO QUADRIO CURZIO (Università Cattolica, Milano) - ROBERTO RUOZI (Università Bocconi) - ALDO SCOTTO (Università di Genova) - ROBERT M. SOLOW (Massachusetts Institute of Technology) - SERGIO STEVE (Università di Roma) - MARIO TALAMONA (Università di Milano) - SHIGETO TSURU (Hitotsubashi University) - BASIL S. YAMEY (London School of Economics and Political Science).

DIRETTORE - EDITOR

dal 1954 al 1983 (from 1954 to 1983): TULLIO BAGIOTTI
dal 1983 (from 1983): ALDO MONTESANO (Università di Milano)

Segretaria di Redazione (Editorial Secretary): ANNA BAGIOTTI CRAVERI

RIVISTA INTERNAZIONALE DI SCIENZE ECONOMICHE E COMMERCIALI (INTERNATIONAL REVIEW OF ECONOMICS AND BUSINESS)

Pubblicazione mensile (A monthly journal). Direzione e Redazione (Editorial Office): Via Teulì 1, 20136 Milano (Italy), Tel. 8399031, C.c. postale 47300207. Editrice (Publisher) Cedam, Padova, C.c. postale 205351.

Abbonamento annuale 1987 (Subscription rate 1987): Italia (Italy), Lire 100.000; estero (abroad), Lire 125.000. Collezione completa, 1954-1986, disponibile (Whole set of back issues, 1954-1986, available).

CONDIZIONI DI ABBONAMENTO AI PERIODICI «CEDAM»

L'abbonamento è annuale e data dal primo numero dell'anno. Il rinnovo è tacito se non disdetto entro il novembre con lettera raccomandata. La semplice reiezione dei fascicoli non viene considerata disdetta. Il canone s'intende anticipato (contrariamente può essere interrotto l'invio dei fascicoli) e si versa direttamente alla Cedam di Padova o ai suoi incaricati muniti di speciale delega, i quali sono tenuti a rilasciare ricevuta su moduli recanti il marchio Cedam numerati progressivamente. Ai correntisti con pagamento rateale viene accordato addebito in conto quota abbonamento col 20 per cento di aumento sul canone.

Il foro competente è quello di Padova.

Direttore responsabile: Aldo Montesano - Autorizz. Tribunale di Treviso N. 113 del 22-10-54



Rivista associata all'Unione della Stampa Periodica Italiana

Tip. Leonelli - Villanova di Castenaso (Bo)

Proprietà letteraria - Stampato in Italia - Printed in Italy

THORSTEIN VEBLEN: THE THEORY OF EVOLUTIONARY ECONOMICS AS A SOCIAL SCIENCE

by
ROBERT GRIFFIN *

Abstract

Thorstein B. Veblen (1857-1929), critical originator of institutionalist inquiry in political economy, extends the historical sequence of classical political economy running from Petty to Marx. Veblen disputed the prevailing "preconceptions" of his time, seeking their positive transformation into a scientific social science of political economy. Central to his theory is an architectonic analysis of social relationships with their evolving, institutional differentiations. Veblen's work reflects his effort to synthesize a Marxist view of historical materialism with the newer sciences of psychology and sociology. Institutionalist economic theory has presently gained respect and adherence that is world-wide.

Economic theories are integrally related to the social milieu from which they emerge. They evolve from, reflect, and seek to provide explanations for the historical process in which they participate. While, therefore, economic theories are themselves the products of the socio-economic matrices in point of origin, the most influential political economists in history have always reflected the most advanced principles of science in their work. In this unique role, economists of the classical tradition from Quesnay and Smith to Marx and Veblen, as well as their cultural heirs, have striven to build a science of economic behavior, one that would prove useful both to serving the needs of people to earn a livelihood and to meeting the growing requirements of public policy making. It is in this sense that Thorstein Veblen (1857-1929) critically reviewed the preconceptions of his predeces-

* Southern Connecticut State University, New Haven, Conn. (U.S.A.).

sors and contemporaries in political economy while, at the same time, seeking to offer a constructive approach to continuing the classical heritage of production and social surplus theory. In point of historical priority and scientific sequence, as Garegnani has noted, the classical foundations preceded the century-long dominance of the demand-and-supply approach in political economy¹.

Transferring his own training in drawing sea charts to the problems of political economy during the 1600's, Sir William Petty first formulated the basic concept of economic value: "If a man can bring to *London* an ounce of Silver out of the Earth in Peru, in the same time that he can produce a bushel of Corn, then one is the natural price of the other..."². Qualifying his discovery of "natural price" with the sensitivity and insight of an architect, Petty added: "This, I say, to be the foundation of equallizing and ballancing of values; yet in the superstructures and practices hereupon, I confess there is much variety, and intricacy..."³. In respect to the world of science of the time, Petty's discoveries in political economy rank with the mathematical and physical advances made by Descartes, Newton and Leibniz and with the architectural reconstruction of London by Sir Christopher Wren following the Great Fire of 1666.

It is also noteworthy that Petty completed his *Political Arithmetic*, published seven years later in 1683, as the first work in social analysis to be used on the statistical method of inquiry. However, it was another century and a half before Thomas Malthus founded the first professional society for the study of statistics. By the end of the nineteenth century the spread of statistical analysis had revolutionized Western patterns of reasoning in logic, the natural and physical sciences, and the new world of social science. Thorstein Veblen himself urged the adoption of statistical concepts and methods for inquiry into socio-economic development as first outlined by Petty. Ironically, Thorstein Veblen made very little use of statistics in his published works on the formal structure and history of modern business and industry. He was very influential, however, in having his qualitative studies pursued quantitatively by Wesley C. Mitchell, particularly the empirical investigations of Mitchell done on the money economy and the historical role of credit in the recurrence of business cycles. The quantitative researches of Mitchell, derived to a great extent from the qualitative inquiries of Veblen, have come to be generally recognized as monumental classics of political economy in the first half of the twentieth century.

¹ GAREGNANI (1984, p. 292).

² ANIKIN (1979, p. 64).

³ *Ibid.*, p. 64.

Thorstein Veblen: The Origin and Structure of His Economic Theory

Born in the second half of the nineteenth century, Veblen spent a rather isolated and obscure childhood and youth as the intellectually talented son of Norwegian immigrants who lived in the rural, mid-western states of Wisconsin and Minnesota. He pursued graduate studies in philosophy at Johns Hopkins University, where he learned the logic of relatives from Charles S. Peirce, and received his doctorate in philosophy from Yale University where he devoted himself to the philosophies of Kant and Hegel. Periodically unemployed, except for ten years of teaching at the University of Chicago and briefer episodes at Stanford University, the University of Missouri and the New School for Social Research, Veblen wrote prolifically. Often referred to as the American Marx, although he exhibited none of Marx's ability for political organization, Veblen published ten books and numerous articles and reviews in professional journals and periodicals. He died in 1929, slightly more than a month before the "crash" he foresaw.

In his role of critic, Veblen commented caustically on what he regarded as the pre-scientific modes of reasoning and inquiry which then prevailed in the dominant schools of political economy. He sought to point out the serious limitations of the "taxonomic" and equilibrium analyses of economic issues, while acknowledging the importance of their contributions to the history of economic thought. He opposed the hedonistic analyses which were derived from the older associationist psychology of the British empiricist school from Hume to Bentham and the Mills. To the German Historical School Veblen voiced his disapproval of their treatment of history as little more than data collection without much regard for inductive generalization and causal connection. Favoring empirical approaches to the study of social science, Veblen rejected religious and metaphysical notions of divine purpose in the universe and in human affairs. Teleology, he argued, had no place in the considerations of modern science. Hypothesized archetypes of "economic man" and abstracted tales of Robinson Crusoe with his "Man Friday" he rejected as superficial substitutes for the real work of explaining the origins and growth of our economic institutions.

As alternatives to the preceding criticisms of the dominant preconceptions of political economy at the turn of the century, Veblen's own position favored treating political economy as an evolutionary social science, one that would be based on the most recent principles and methods of modern science and "the logic of mechanical effect" (the relational reasoning systems of C.S. Peirce and William James) in modern industry and business enterprise. Differing from the traditional, static approaches that concentrated on

problems of economic equilibrium, Veblen proposed the investigation of *dynamic change* in science and industry that provided the leading edge for economic development. With respect to the newer, scientific developments of psychology introduced by the monumental work of William James in 1890, Veblen advocated replacing the associationist account of hedonistic, economic man as a "lightning calculator of pleasures and pains"⁴ with the direct observation and statistical confirmation of human behavior, propensities and economic habits being developed by James, Dewey and McDougall. As a corrective to the German Historical School, Veblen explained that historical research of a scientific nature called for inquiry into sequential processes that could demonstrate a line of cumulative causation and the evolution of economic institutions as "on-going concerns" running on into the future. This latter approach, it seems, has been more widely favored in the United States during the last hundred years as a more realistic approach to the development of regulatory policies at the administrative and management levels. In place of speculations about divine teleology, Veblen recommended the scientific study of purposive conduct in human society. This recommendation represents an American-Kantian transformation of teleological physics into the physical teleology of human experience and society. The romantic Robinsonades which have served little purpose other than to illustrate and reinforce the simplest beliefs of early capitalist society and colonialist prejudices, according to Thorstein Veblen, needed to give place to the modern, scientific inquiry into political economy. In line with the architectonic outline first sketched by Sir William Petty, Veblen emphasized that any accurate accounting of the actual ways in which we develop our modern industrial society must provide an adequate explanation of the habit, custom and institutional differentiations that have led to the more general super structure of our social existence⁵. In this respect, the evolutionary economics of Thorstein Veblen reflects a theoretical elaboration of the Marxist idea of historical materialism as set forth by Marx, Engels, Bebel and Labriola. Veblen, however, makes extensive use of William James' theory of habit which was relatively unfamiliar to his Marxist contemporaries.

"The Darwinian controversy is, in large part, a question of logic", asserted C.S. Peirce in 1877. "Mr. Darwin proposed to apply the statistical method to biology"⁶. Likewise, Veblen the student of Peirce and Mitchell the student of Veblen looked to "the logical basis of economic theory"⁷.

⁴ MITCHELL W.C. (1947, p. 144-45). Cf. BRONFENBRENNER (1985, pp. 19-20).

⁵ *Ibid.*, pp. 41-145.

⁶ PEIRCE (1960, 5.364).

⁷ MITCHELL L.S. (1953, p. 234. Letter to Derrick Lehmer, dated October 23, 1912).

The logic of modern machine technology, stated Veblen, is "a logic of masses, velocities, strains and thrusts". This mode of reasoning stands somewhat in contrast to the pre-industrial artisan's thinking in terms of "personal dexterity, tact, training, and routine"⁸. This newer "logic of mechanical effects" also contrasts with the culturally inherited tendency toward a teleological bias. In the world of accounting, for example, "The logic and concepts of accountancy are wholly impersonal and dispassionate; and whether men's use of its logic and concepts takes the elaborate form of a set of books or the looser fashion of an habitual rating of gains, losses, income, and outgo in terms of price, its effect is unavoidably in some degree to induce a statistical habit of mind".

With respect to the history of ideas in economic development, Veblen again reflects a current of conceptual history in which C.S. Peirce was a leading figure. Characteristically, Thorstein Veblen saw three stages in the history of reasoning since the medieval era: 1. Recourse to authority, whether that of the revealed word of the Heavenly King or that of the potencies and qualities ascribed to the hierarchical dignitaries under the finished, earthly scheme of feudalism. The discipline of handicraft, however, trained the general population to push its inquiries toward the liquidation of the First Cause. 2. "In the later age the good of man gradually, insensibly supplants the glory of God as the ultimate ground of systematisation"⁹. Recourse in this period of "enlightenment" and "humanism" is to an *a priori* Order of Nature, 'First Principles', and an accompanying scheme of Natural Law. It was a world in which Adam Smith and the Physiocrats could both feel "God's in his heaven, and all's right with the world". 3. The relational logic of modern science begins to appear and make its ascendancy with the impact of the industrial revolution. Under its standards of scientific hypothesis, accountancy and cumulative causation, "what is alone materially competent in the modern scientific apprehension is such an explanation as will make things plain in terms of matter and motion, working a change in the constitution of things by displacement through contact and pressure"¹⁰.

Political economy, in the thinking of Thorstein Veblen, might have been defined as the social science of production, distribution, exchange and consumption interrelationships of human behavior; the superstructures which are produced from social relationships and which evolve as socio-economic institutions are for the most part determined by the relatively constant psychological factor of "habit". Habit, as analyzed by Peirce and

⁸ VEBLEN (1964, pp. 241-45).

⁹ *Ibid.*, pp. 257-58.

¹⁰ *Ibid.*, p. 264.

James, is the logical relation between sense-perception and practice. To study our economic habits and customs, therefore, is to inquire into the logical basis of economic theory.

In the classical tradition of Petty, Quesnay, Smith and Marx, Thorstein Veblen makes human activity, science and industrial production the starting point for his theory of political economy. Then, carefully generalizing from the ethical and educational significance of James' theory of habit¹¹, Veblen formulates a set of ideas for analyzing the base and superstructure of socio-economic development. Although he never published a formal statement of his system, the evolutionary theory of economic institutions which runs through all of his major books and articles may be outlined as follows:

- I. Human Activity and Social Existence
 - A. Purposive Behavior
 - B. Evolved Propensities of Human Conduct
 1. Instinct of Workmanship
 2. Parental Bent
 3. Idle Curiosity
- II. Habit Formation as the Base of Socio-economic Superstructure
- III. Customs as Generalized Habits Building Social Structure
- IV. Institutions as Generalized Complexes of Habits and Customs
 - A. Institutional Standards, Norms, Codes, Rules
 - B. Collective Habits of Thinking in Logic, Communication and Culture
 - C. Social Change
 - D. Institutional Development and Social Evolution
 - E. Economic Institutions and Systems as Products of History
- V. Community as the Ensemble Product of Interrelated Institutions
- VI. Spiritual Attitude, Scheme, or Prevalent Theory of Life
- VII. History as Cumulative Causation and Cumulative Development
- VIII. Ideal of a Cosmopolitan republic of Knowledge and Intelligence.

Let us now take a larger, more detailed look at the Veblen architectonic of political economy.

¹¹ JAMES (1948, p. 134). James' sequence of maxims, now widely adopted for educational and therapy programs, proceeds as follows: 1. "Make our nervous systems our allies instead of our enemies". 2. "Continue to repeat the pattern... *Continuity of learning* is a great aid to the nervous system". 3. "Build associations..." 4. "Find some way to attend to the habit favorably each day and use some effort to encourage its growth through exercise".

I. *Human Activity and Social Existence*. Human activity is the basis for socio-economic evolution. *Economic activity* and productive industry, therefore, becomes the main focal point for "An Economic Study of Institutions" ¹².

The life of people in society, according to Veblen, is a struggle for existence. Human beings are more or less aware of being centers of unfolding, impulsive activity which is purposive. The notions of the Physiocrats and Adam Smith about a teleological order of nature are treated by him as misdirected, metaphysical speculation which realistically concerned the purposive, teleological direction of human experience ¹³.

The forces which shape human evolution and social structure he classifies as partly human, partly non-human, and a human subject with a definable physical and intellectual make-up ¹⁴. The resulting social existence is a dialectic of interacting forces serving individual and social need-hierarchies.

Human activity, viewed as an evolutionary differentiation from the activity of the other species, has three unique features, or "instincts": 1. the instinct of workmanship 2. the parental bent 3. the instinct of idle curiosity ¹⁵. Disclaiming any intent to reduce these propensities, or evolutionary biases of the human race to "isolable tropismatic sensibility" or single stimuli of a visceral nature ¹⁶, Thorstein Veblen considered the triumvirate of instincts as significant propensities which have come to characterize the purposive behavior of the human species throughout its history, especially in its struggle to survive and save itself alive from catastrophe at various times in history. "All instinctive action is intelligent and teleological" ¹⁷.

The instinct of workmanship expresses itself in the production of "practical expedients, ways and means, devices and contrivances of efficiency and economy, proficiency, creative work and technological mastery of facts". Functionally, "the instinct of workmanship is a proclivity for taking pains" ¹⁸. Additionally, Veblen ascribes a decisive role to labor's workmanship-like propensity for instituting every major advance in history. Moreover,

¹² Subtitle to *The Theory of the Leisure Class* (VEBLEN 1953). Unless otherwise indicated, references to Veblen's evolutionary analysis of socio-economic institutions are those I have selected from his basic theory first stated in *The Theory of the Leisure Class*.

¹³ *Ibid.*, pp. 29, 83, 132-40.

¹⁴ Cf. ROUSSEAU, *Emile*, Bk. I, "The Meaning of Education".

¹⁵ JAMES (1950, pp. 410, 430, 627-628). McDUGALL, (1905, chps., I, III, X). The Veblen triad of instincts actually parallels the ancient Greek emphasis in rhetoric on the *ethos*, *pathos* and *logos* principles of persuasion in Aristotle's *Rhetoric*.

¹⁶ VEBLEN (1953, p. 27).

¹⁷ *Ibid.*, pp. 31-32.

¹⁸ *Ibid.*, p. 33.

it is particularly significant that Veblen defines the workmanship instinct as "serviceability for the ends of life, whatever these ends may be;..."¹⁹. Since Thorstein Veblen also gives an accounting explanation of Marx's theory of surplus value as the excess of serviceability over cost, his connection of surplus value with serviceability and workmanship is quite on a par with Marx's delineation of the labor theory of value and its relation to socially necessary production.

While the sense of workmanship is the primary disposition which conduces "directly to the well-being of the race, integrally bound up with it is the parental bent"²⁰. The "parental bent" encompasses much more than the procreation of children. It accounts most, according to Veblen, for the rejection of race-suicide by all peoples, extending also to persons who show no evident interest in the procreation and rearing of children. Likewise, there is a general disapproval of every suggestion that the present generation should wilfully make life more difficult for the next generation. Collectively, the parental bent finds its social expression in "the sentimental concern" for the life, comfort and future welfare of the community in general. It lends sentimental reinforcement to the efforts toward economy and efficiency for the common good, causing general disapproval of economic waste. Arguably, the spirit of economy can be considered a function of the instinct of workmanship, in which case the point is that workmanship is strongly sustained by a certain logic of sentiment exhibited "by a parental solicitude for the common good"²¹.

"Idle curiosity" is the propensity which impels people to "want to

¹⁹ *Ibid.*, p. 31. RUTHERFORD (1984) provides a useful analysis of the Veblenian outlook on cumulative sequence in relation to the cycle of change in which institutions affect technology and technology affects institutions (pp. 343-44). But, his argument that Veblen required the "instincts" as exogenous factors "to start the analysis" (pp. 332-33) is off the mark. Veblen, like Marx, does not require any exogenous First Cause or Final Telos. For Veblen the argument begins with movement and social activity and ends with the open-ended movement of the world economic community as an "on-going concern" running on into the future. With Veblen as with Goethe, "Im Anfang war die Tat" (*Faust*, Scene III). The "instincts", therefore, are secondary and *endogenous* to the larger field of industrial activity, as vectoral forces in the evolution and history of man's struggle to survive.

Similarly, technology, as described by Rutherford, is not at all the primary cause of insitutional change. Institutions being of the nature of habit, argues Veblen, are *inherently* subject to change as part of the historical process of science and industry. "The technological advance is due to the engineers and their sense of workmanship". DENTE (1977, pp. 126-65).

²⁰ VEBLEN (1964, pp. 25-27).

²¹ PEIRCE (1960, 5.357): "We are in the condition of a man in a life and death struggle; ... the only assumption... is the hope of success. So this sentiment is rigidly demanded by logic".

know things, when graver interests do not engross their attention" ²². It complements the instinct of workmanship by accelerating gains in technological insight, as well as by disturbing institutionalized knowledge through doubt and inquiry. It is "idle" only in the sense that elements of utilitarian advantage do not really affect its habitual exercise, although the knowledge gained by it may eventually serve the ends of workmanship. Considered in the light of its long-term consequences, idle curiosity, assisted by talented individuals in whom it runs to a higher degree, is historically significant by means of cumulative causation for "the most substantial achievement of the race, — its systematised knowledge and quasi-knowledge of things".

II. *Habit Formation*. The *base* of any politico-economic superstructure, according to Veblen, is the complex of social relations and, so to speak, the fly-wheel of industrial habits that serves to give a dynamic balance to the input-output flow in the social mode of existence and production. Habits, formed as unfolding pathways of human practice, constitute the major social force which gives direction and purpose to human expression. For example, with the rise of modern capitalist organization, the ritual habit of making money has come to serve a two-fold purpose: it both facilitates and obstructs the daily task of earning a living. In the life of the modern community there are, consequently, two habitual lines of economic activities: one being the industrial production of raw materials into serviceable goods; the other being the series of disruptive bargains and exchanges by which people determine for themselves the amount each individual will take to himself or herself and the amount the others will have. The material welfare of the community largely depends on the first set of habitual activities — those of human production, distribution and consumption. The habitual exchange activities — price, profit, accounting, banking, finance, insurance and, by extension, modern credit card systems, etc. — contribute to the welfare of the community only in so far as they enlarge the quantity or improve the quality of the goods and services which result from the original expenditure of energy and material resources. Paradoxically, the routine of money-making is a means of producing goods and services; but, under the rule of business enterprise in modern history, the priority of the social relationships has become completely upset. Under the guardianship of the capitalists, goods and services are now produced to make money. From this historically new primacy of the pecuniary economy there flow in seemingly endless

²² VEBLÉN (1964, pp. 85-89); ROUSSEAU (*Emile*): "The Third Stage of Childhood". "There is a quest for knowledge which comes only from the desire for reputation as a scholar; another which springs from natural curiosity about anything interesting near and far".

sequences whole multitudes of strange distortions that have their origins in the faulty habit formations cultivated through the institutional forms of capitalist society. These distortions, in the estimation of Veblen, are particularly ill-adapted to the scientifico-technological requirements demanded by the routines of modern industrial society²³.

At any given point of socio-economic development, the institutionalized "standard of living" is also of the nature of habit. It is comprised of those habitual norms of occupation and consumption which people come to regard as most important for their livability in the community at any one time. Logically and statistically, a standard of living can be analyzed into its component habits and be measured by a rating scale with a corresponding technique for registering responses to given stimuli within a definable population.

At the level of subsistence, Thorstein Veblen classifies the fundamental hierarchy of habits into two types. These types, in turn, correlate with the two main levels in the hierarchy of human needs.

1. Habits of existence. Survival, maintenance and reproduction represent the most ancient and engrained ensemble of habits that govern an individual's life in society. For this reason, the basic habits of existence tend to be the most persistent and imperative. For the same reason, historically and anthropologically the oldest habits are those associated with the earliest communal work of hunting, fishing and agriculture.

2. Beyond the routines geared to the basic necessities of existence and social reproduction are the habits stemming from the "higher wants". These habits, also, come to form a hierarchy in a somewhat irregular and often variable gradation; e.g., the use and cults of drugs, a "need of salvation" religiously, or the social need for status and to be "of good repute". This spiritual hierarchy of socio-cultural wants and habit formation also helps to explain the strenuous efforts of many people to attain "a conventional standard of decency", together with a socially sanctioned level of "conspicuous consumption".

Continuity and persistence of habits cause them to win an appearance of being securely rooted in the life of the community. Paraphrasing William James, Veblen states: "In general, the longer the habituation, the more unbroken the habit, and the more nearly it coincides with the previous

²³ VEBLER (1953, p. 83). MITCHELL W.C. (1936, Introduction x1-x1vii). This concept of habit formation is derived *via* James' *Psychology* from Charles Darwin's treatise of 1872 on the expression of emotions and mental activity in humans and animals, published also in Stuttgart and St. Petersburg in 1872, from where it also found its way into the hands of the Marxist G.W. Plechanow. Cf. PLECHANOW (1955, p. 55).

habitual forms of the life process, the more persistently will the given habit assert itself"²⁴. To illustrate his point, Veblen cites the rise of the handicraft system of the Late Medieval, Renaissance and national dynasty periods. While admitting the relative nature of socioeconomic systems, of politico-economic ideologies, and of differing national development, he selects the rise of handicraft as the point of origin for the historical evolution of the modern era²⁵. Under the disturbing, reconstructing influence of the instinct of workmanship, the dialectical processes of social history led to the decay of feudalism with its "servile agricultural organisation of industry", on the one hand, and, on the other hand, to the rise and habitual outlook of the handicraft system with a two-fold character: "technological and pecuniary".

Between the 1200's and the 1600's there grew up in Europe a handicraft routine of everyday life, particularly in the towns, that was characterized by a spirit of self-help combined with pride in workmanship. To maintain their privileges in the towns, the craftsmen organized into guilds. The institutionalized guilds, in turn, began to set down craft rules for regulating and governing the output of their goods, chiefly for the purpose of maintaining a price that would secure them an adequate livelihood. Veblen viewed the artisans, therefore, as the precursors of the modern engineers, industrially, and the "efficiency engineers" among the twentieth century capitalists who calculate efficiency "proximately in terms of mechanical performance, ultimately in terms of price, and more particularly in terms of net gain". Out of a meticulous regard to minute details, the craft guildsmen gave rise to the doctrine of "natural rights" of property and to "business principles" that later came to prevail anachronistically in the era of machine industry and financial investment.

By means of a "division of labor", states Veblen, the craft system became increasingly differentiated into industry and business. The latter came to specialize in ownership, property rights, monetary exchange and the building of associations to protect and promote the business community. This evolution of "the business community" was parallel, although increasingly irrelevant to the continued development of the industrial system and of modern science and technology.

As is dramatically evident in the history of operatic productions created by Giovanni and Baron Ludovico Burnacini, his son, in the late 1600's to the grandiloquent constructions of the Galli-Bibiena family in the 1700's by means of a more elaborate division of labor, masterworkmen with

²⁴ VEBLEN (1953, p. 83).

²⁵ VEBLEN (1964, pp. 209-215).

means came to employ other workers without means. The members of a craft guild who were fortunate enough to continue at their trade, as in the case of the Galli-Bibiena's, grew quite wealthy with the increasing volume of trade in Europe and sometimes even took on titles of nobility²⁶. In the case of the textile industry, wealthy craftsmen would organize both cottages and factories into sweatshop industry in order to meet the increasing demands of the market; or, they would engage in overseas trade as jobbers conducting their transactions over a longer term, thereby withdrawing farther from the actual production line in the craft. In the 1500's, the Fugger banking family had already grown into a financial dynasty that maintained only distant and impersonal relationship to their craft industries.

Continuity, persistence and elaboration of pecuniary habits also marked the commercial practices of the business world in its early history. With the advent of double-entry bookkeeping practices among the Florentines of the Renaissance, new patterns of rationality began to be drilled into the minds of tradesmen. Through the impact of modern accounting practices there arose a demand for more standardization of the monetary unit which, in turn, provided a common denominator through which all values could be expressed: census figures, employment and unemployment statistics, occupation and work categories, wage and salary rates, etc. This unbroken, relatively new habit of accounting has persisted historically from its origin in the Renaissance to a modern era that Veblen might well have anticipated, the era of computer science and technology. Not only has the introduction of accounting practices drilled further habits of rationality into the thinking of people in the western world; it has, through the added modifications of feedback mechanisms in our own time, reinforced and elaborated our social patterns of reasoning to the point where presently the new environment of wages and prices forces us to plan on pain of survival.

III. *Social Structure*. Above the basic level of existence with its social mode of production, the architecture of society is made up of the habitual effects which then become generalized the customs and institutions by which the community organizes and conducts its cultural practices.

IV. Institutions, as generalized complexes of habits and customs, are the structural components of the community. As such, they are the habitual standards, "canons of conduct", for carrying on the life processes of society. To the classification of institutions also belong the norms and codes of

²⁶ GRIFFIN (1972).

conduct, written or unwritten, verbal and nonverbal, by which a society governs its principal mode of existence. Such standards have been found in all societies, whether they be the hunting and fishing codes of the Eskimoes, the currently sanctioned conduct of the Zuñi Indians of the Southwestern United States and their agricultural practices, or the vast industrial and financial ensembles of modern, scientifico-technical living which are shaping the architecture of the modern world. Under the rule of capitalist enterprise, observes Veblen, the canons of conspicuous consumption and honorific waste continue to hold sway. With the dispossession of the old order and the introduction of a new, industrial order based on economy and efficiency in the service of the community, there could be a rational organization of society with full employment and an avoidance of waste and duplication of work²⁷. The transition to a rational industrial republic, however, must be organized and led by a technical elite, General Staff of technicians, engineers and scientists²⁸.

Additional canons of conduct which Veblen found to prevail among the luxury loving members of the leisure class in his day were the canon of "pecuniary emulation" and that of "conspicuous leisure". The habit of pecuniary emulation originates in the industrial efficiency of an ownership economy where a consistent system of property in goods has been installed. At a luxury level of production, the leisure class is actually not interested in the consumptive purposes for accumulating goods. Since the main motive which lies at the root of ownership is emulation, this same motive continues to govern the development of the economy at the luxury level. Luxuries are acquired for the honorific evidence they display of the owner's "prepotence". Luxury in the form of money likewise confers honor. "It is an invidious distinction". As a generalised social norm, relative personal and family success, shown by comparison to exceed the wealth displayed by others, is the practice of pecuniary emulation which has become the conventional purpose of living in an advanced capitalist society. Pecuniary emulation being the institutional norm of primary importance to the monied leisure class, chief among the secondary demands of emulation is an obvious abstention from productive work. To this end the ancients instituted the practice of slavery and saw their wealth to consist in the number of slaves held for

²⁷ VEBLEN (1965, pp. 162-169).

²⁸ DENTE (1977, pp. 165-181). In spite of repeated refutations, the specious connection of Veblen with the Technocracy, Inc., aberrations of Howard Scott still turn up occasionally: PIROU (1946, pp. 89-105); GROSS (1980, p. 80): "Veblen's vision was later embodied in Technocracy, Inc., ...". For Veblen technician was a synonym for "worker"; his class of technicians, engineers and scientists a 1920 equivalent of Marx's working class.

personal service. At the turn of the century, Veblen registered the educational evidence of conspicuous leisure as: "knowledge of the dead languages and the occult sciences; of correct spelling; of syntax and prosody; of the various forms of domestic music and other household art; of the latest proprieties of dress, furniture, and equipage; of games, sports and fancy-bred animals, such as dogs and race-horses"²⁹. Since the code of proprieties itself shows an incompatibility with the productive work of the general population, Thorstein Veblen maintains that the code of decorum is a form of economic waste that serves as a "canon of conduct for the classes beneath"³⁰.

Because the institutions are collective habits of thinking, they also prevail as the dominant patterns of reasoning and make up the social substance for the concrete, individual relationships in the community. This social substance comes to be expressed as modes and fashions of dress, manners, family structure, property laws, religious dogmas, social class distinctions, etc. Also included as institutionalized communication ritual is the business man's routine verbalizations as to what are the proper limits, management prerogatives and ownership responsibilities. Dialectically, these business verbalizations function in opposition to the frustrated, exploited sentiments that stem from the bias of workmanship, which under the wage system continues to impel the workmen to fulfill their sense of duty to the discipline of the machine process.

As institutions are the products of human activity and social process, however, so they are themselves subject to *change*. Institutions change with the changing circumstances and changing requirements of history. Consequently, any fancied return to the age of national dynasties such as are often found in the maudlin tales of a bygone era or featured in journalistic weeklies is quite literally the fiction of gothic romance, ill-adapted to a world that is daily becoming more attuned to the sounds of industrial processes in a scientifico-technological age. "It seems possible to say this much", stated Veblen in 1903 and with characteristically uncanny foresight, "that the full dominion of business enterprise is necessarily a transitory dominion. It stands to lose in the end whether the one or the other of the two divergent cultural tendencies wins, because it is incompatible with the ascendancy of either"³¹. The complexity and the duration of the transitional

²⁹ VEBLEN (1953, pp. 33-51). "In this view the university is conceived as a business house dealing in merchantable knowledge, placed under the governing hand of a captain of erudition". VEBLEN (1962, p. 62).

³⁰ *Ibid.*, p. 51.

³¹ VEBLEN (1958, Chp. X, "The Natural Decay of Business Enterprise", p. 189).

process, however, are far greater and much lengthier than even Veblen projected for his "calculable future".

Institutional development is the equivalent of social development. It is, in fact, the evolution of society. For this reason, the impact of the industrial process and its logic of mechanical effects on social living is much more than the aggregate production of gadgets and appliances. On the logical side, C.S. Peirce expressed the enginery of modern thought in a maxim for the "Logic of Science":

Consider what effects, which might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object ³².

With reference to political economy, Thorstein Veblen saw the idea of "tha machine process" evolving its purpose of building an industrial republic that would include new work relationships for "the civil engineer, the mechanical engineer, the navigator, the mining expert, the industrial chemist and minerologist, the electrician,..." ³³. The evolving system of modern scientific-technology would also include "the work of the inventor who devises the appliances of the process and that of the mechanic who puts the inventions into effect and oversees their working". This position does not mean that Veblen's theory becomes "micro-economics". Rather does it mean that a realistic assessment must be made of the whole evolutionary process of economics, analysing the infrastructural development from the micro- to the macro-economic level ³⁴.

In his *Theory of Business Enterprise* Veblen proposes that *capital* be viewed *prospectively*, as businessmen actually do, and not only *retrospectively*, as economists have frequently done. Taken retrospectively, as is mainly the case with Quesnay and Marx, capital is explained as the historical accumulation of labor- and surplus-value, the emphasis tending to fall on capital as the product of past history. Veblen accepts this notion of capital, interpreting it as a cumulative fund in the current processes of industry and business, and draws attention realistically to the prospective businessman's view of "venture capital", or capital turned to account and increased by marketable credit instruments which, in turn, may also appear as collateral for a further expansion of credit. Both theories, argues Veblen, need to be

³² PEIRCE, *op. cit.*, 5.402.

³³ VEBLEN (1958, Chp. 1, p. 9).

³⁴ TODARO (1981, p. 19, 21): "Why is it important to view economies as social systems and to go beyond simple economic factors in analyzing development problems?"

employed for a full account of what happens in the way of social and institutional development under capitalism.

In the case of business organization, a start is commonly made from a substantial body of immaterial goods; e.g., a franchise and "good-will", or the control of a natural resource such as water, gas, or even air; or, a special process, secret or patent; or a combination of intangible assets. Capitalizing on these intangible assets, a private company or partnership may convert to a corporation, unless the firm began its existence as a corporation. In the modern business world, the owners of the corporation ordinarily choose to absent themselves from responsibility for the running of the concern. They turn the direction of the business over to a management group that promotes the "good-will" base of the firm by means of trade-marks, clientele, and such financial relations as will position the company for assuming a monopoly status, either locally or over a wider area of trade and finance. Indeed, from the time of Veblen to the present era of conglomerate mergers, the humble, institutional origins of business expansion have taken on the form of international conglomerates through merger movements and purchase of stock. In contrast to the evolution of the industrial process through its logic of mechanical effects, guided by the basic instincts – the instinct of workmanship, the parental bent, and the instinct of idle curiosity – the evolution of the business corporation and its more elaborate, conglomerate giants on the contemporary world scene have their foundations and social substance to a greater extent than is often realized in the corporate, immaterial assets.

The Veblen theory of evolutionary economics also has implications for the history and theory of political economy in general. Since all economic institutions are the products of historical development, the vocabulary terms of political economy – wage, price, profit, property, insurance, credit, industry, banking, etc. – can also be analyzed for their institutional significance. Since all of the terms mentioned and many more can be analyzed historically and relationally as the economic ensemble-products of socio-economic evolution, what comes to be at issue frequently is not just a "price", but the whole price-system; not an isolated worker's "wage", but the wage-system; not the turn of a "profit", but the entire profit-system; and so on to an examination of the ownership property system, the insurance system, the credit system, the systems of science and industrial technology, the banking system, etc. In *The Theory of Business Enterprise*, therefore, Thorstein Veblen takes issue with the capitalist system. According to his projections, the further evolutionary resolutions of capitalism would take the direction of either the dynastic state or of a socialist industrial system. The funda-

mental instincts of humanity favored the latter, but history has no final term.

V-VI. The social architectonic reaches a completion in *the spiritual attitude, scheme, or prevalent theory of life*. According to Veblen, the community's spiritual attitude is projected from its base in the life processes of society. According to Veblen, the community's spiritual attitude is projected from its base in the life processes of society. This objectification, communicated verbally and non-verbally through the social structure by society's members, produces the dominant social theory. The preconceptions, also economic preconceptions, of the ruling-class scheme historically are of a much older vintage, having a durability and continuity which tends to run on into the future. This element of cultural inertia makes them more resistant to change, less susceptible to alteration. For this reason, the robber barons of the late 1800's are all long since dead; yet, their financial empires and spiritual attitude of imperialism have continued to serve the absentee-ownership interests of their heirs through several wars in the present century.

VII. *Human history* Veblen views as a major example of cumulative causation and cumulative growth. Through a variety of socio-economic stages: the barbarian stage, ancient slave society, feudalism, the age of the landed gentry and baronic nobility, the rise of craft guilds with their claims of natural order and natural rights, industrial capitalism of the 1800's, and the present stage of international socialism and the corporate state of financial capitalism — human history reflects a series of class struggles. These class struggles have also entailed the cumulative impact of the inherited institutions possessing an essentially predatory nature, on the one hand, and of those which have had their origins in workmanship, husbandry and industry, on the other. As mentioned, history has no final term; the direction and adaption of humanity's needs in relation to its cultural heritage and future goals is still undecided.

VIII. The Cosmopolitan Republic of Knowledge and Intelligence was one of the highest ideals in the system of Veblen. It inspired him to be an internationalist. Like all persons concerned with world peace and justice, Thorstein Veblen hoped for the establishment of an international, anti-imperialist community of nations. He rejected the original establishment of the League of Nations in the early 1920's, regarding it then as a virtual league of thieves opposed to the long-term, peaceful interests of the human race.

His own Cosmopolitan Republic ideal bears a more kindred spirit to the present organization and work of the United Nations.

During the last decade there has been a veritable renaissance in the study of Thorstein Veblen. Joan Robinson, proclaiming "The Abdication of Neo-classical Economics"³⁵, drew attention to Veblen's broader interpretation of "capital", and to his emphasis on the role of science and technological knowledge in the development of economic institutions. Gunnar Myrdal, doing a *mea culpa* at the University of Wisconsin in 1977, described his pilgrimage from being a zealous, narrow-minded initiator of econometrics to his later appreciation of economics as a social science to be seen in historical perspective. He views the Veblen inspired study of economic institutions as gaining ground "because a broader approach will be needed for dealing in an effective way with the practical and political problems that are now towering and threatening to overwhelm us"³⁶.

In 1977, Leonard Dente completed and published a doctoral dissertation at the New York University on *Veblen's Theory of Social Change*. It included a rare exposition on the unity of Veblen's economic theory. This very fine work, however, fell into relative obscurity until recently³⁷.

More recently, Marxist-Leninist scholar Herbert Aptheker gave extended and conventionally guarded praise to Thorstein Veblen for his supportive recommendation of Karl Marx to social scientists at a time when to do so was tantamount to sacrificing one's academic career³⁸.

In the most recent issue of *The American Economic Review* for December, 1985, articles by William J. Baumol and Martin Bronfenbrenner draw attention to the growing importance of Veblen's concepts and the critical insights of the institutionalists³⁹.

Two of the recognized books, bearing the stamp of the Veblen influence, are *History of Economic Thought: A Critical Perspective* by E.K. Hunt, and *The Worldly Philosophers* by Robert L. Heilbroner. Of Veblen's influence Hunt writes: "I believe, with Veblen and Marx, that capitalism is not the highest stage of human development and that ... human beings ... will open a vista of passionate possibilities hardly dreamed of during the reign of capitalism"⁴⁰. The influence of C.S. Peirce and Thorstein Veblen on Heil-

³⁵ ROBINSON (1980, pp. 37-38, 39-40).

³⁶ MYRDAL, "Institutional Economics", lecture at the University of Wisconsin, Dec. 15, 1977; rep. MYRDAL (1979, p. 106).

³⁷ DENTE (1977); GRIFFIN (1982, pp. 72-76).

³⁸ APTHEKER (1983, p. 19).

³⁹ BAUMOL (1985); BRONFENBRENNER (1985).

⁴⁰ HUNT (1979, p. 458); HEILBRONER (1980).

broner is best felt in the historical development account of Chapter 1, entitled "The Economic Revolution".

The ideas of Veblen and other institutionalist economists have now taken root in Japan, also. Shigeto Tsuru, to whom I am indebted for the biographical notes on Gunnar Myrdal, has made an extensive survey of several contemporary writers and lecturers on institutionalist economics. He likewise recognizes the seminal influence of Veblen. His own lectures, delivered at the University of Milan in May, 1985, are as yet unedited; but they convey the definite impression that Veblen and institutionalist political economy are alive and well in Japan.

The preceding prolegomena to the principles and structure of Veblen's theory of political economy makes no pretense at being exhaustive. It is only my hope that the prolegomena may prove sufficient to encourage others to pursue in detail the challenge of Thorstein Veblen to build a realistic, evolutionary social science of political economy.

REFERENCES

- ANIKIN A.V., *A Science in its Youth: PreMarxian Political Economy*, New York: International Publishers, Inc., 1979.
- APTHEKER H., "Marxism and Social Sciences in the U.S.: A Brief Appraisal", *Political Affairs*, 6/1983, 62.
- BAUMOL W.J., "On Method in U.S. Economics a Century Earlier", *Amer. Econ. Rev.*, Dec. 1985, 75.
- BRONFENBRENNER M., "Early American Leaders – Institutional and Critical Traditions", *American Economic Review*, Dec. 1985, 75.
- DENTE L., *Veblen's Theory of Social Change*, New York: Arno Press, 1977.
- GAREGNANI P., "Value and Distribution in the Classical Economists and Marx", *Oxford Economic Papers*, 1984, 36.
- GRIFFIN R., *High Baroque Culture and Theatre in Vienna*, New York: Humanities Press, 1972.
- , *Thorstein Veblen: Seer of American Socialism*, New Haven, Ct.: Advocate Press, 1982.
- GROSS B., *Friendly Fascism*, New York: M. Evans & Co., 1980.
- HEILBRONER R.L., *The Worldly Philosophers*, 5th ed., New York: Simon & Schuster, 1980.
- HUNT E.K., *History of Economic Thought: A Critical Perspective*, Belmont, Ca.: Wadsworth Publ. Co., 1979.
- JAMES W., *Psychology*, Cleveland, Oh.: Fine Editions Press, 1948 (rep. from the 1892 abridged ed.).

- , *The Principles of Psychology*, II, New York: Dover Publications, 1950 (rep. from Henry Holt & Co., 1890 ed.).
- McDOUGALL Wm., *Introduction to Social Psychology*, Boston, Mass.: Luce & Co., 1905.
- MITCHELL Lucy S., *Two Lives*, New York: Simon & Schuster, 1953.
- MITCHELL W.C., ed., *What Veblen Taught*, New York: Viking Press, 1936.
- MYRDAL G., *Essays and Lectures after 1975*, Kyoto: Keibunsha, 1979.
- PEIRCE C.S., "The Fixation of Belief" (1877), in *Collected Papers of C.S. Peirce*, 5.364, Cambridge, Mass.: The Belknap Press of Harvard University Press, 1960.
- PIROU G., *Les Nouveaux Courants de la Théorie Economique aux Etats-Unis*, Tome I, Paris: Editions Domat-Montchrestien, 1946.
- PLECHANOW G.W., *Kunst und Literatur*, "Briefe ohne Adresse", Erster Brief, Berlin: Dietz Verlag, 1955.
- ROBINSON J., *Collected Economic Papers*, Cambridge, Mass.: MIT Press, 1980.
- ROUSSEAU J.J., *Emile*, Bk. I, "The Meaning of Education".
- RUTHERFORD Malcom, "Thorstein Veblen and the Processes of Institutional Change", *History of Political Economy*, Fall 1984, 16.
- TODARO M.P., *Economic Development in the Third World*, 2nd ed., New York: Longman, 1981.
- VEBLEN Th., *The Theory of the Leisure Class*, New York: Mentor Books, 1953 (rep. from the 1899 ed.).
- , *The Theory of Business Enterprise*, New York: Mentor Books, 1958 (rep. from the 1904 Ch. Scribner's Sons ed.).
- , "The Preconceptions of Economic Science", in *The Place of Science in Modern Civilization*, New York: Russell & Russell, 1961 (rep. from *Quar. Jnl. Econ.* XIII-IV, 1899).
- , *The Higher Learning in America*, New York: Hill and Wang, 1962 (rep. from 1918 ed.).
- , *The Instinct of Workmanship*, (New York: Macmillan, 1914), rep. New York: A. Kelley, 1964.
- , *The Engineers and the Price System*, "A Memorandum on a Practicable Soviet of Technicians", New York: A.M. Kelley, 1965 (rep. from the 1921 B.W. Huebsch ed.).

THORSTEIN VEBLEN: LA TEORIA EVOLUZIONISTICA DELL'ECONOMIA COME SCIENZA SOCIALE

Storicamente Veblen ha proseguito nel solco della tradizione classica che possiamo far partire con W. Petty e che culmina nell'opera di Marx e di J.S. Mill. Di formazione filosofica, con ampi interessi intellettuali e scientifici, Veblen

entra nel campo dell'economia. Il presente lavoro intende mettere in luce come i contributi di Veblen all'economia siano legati a una penetrante rassegna critica delle opere degli economisti del passato e suoi contemporanei all'intento di depurare la disciplina dagli elementi residui di metafisica preindustriale. Sul piano propositivo Veblen offre il suo concetto di una economia evolutiva come metodo di costruzione di un sistema scientifico basato induttivamente, pur nel sostanziale rispetto del quadro tassonomico fissato dai classici. Dai classici Veblen trae l'interesse alla teoria della produzione, mentre fa tesoro della lezione dialettica di Marx e Engels e mostra l'importanza per l'economista della teoria psicologica, specie di Peirce e William James *. I motivi di attualità di Veblen appaiono spesso legati all'importanza che si è disposti ad attribuire alla base "interdisciplinare" dell'economia politica. Ma occorre anche aggiungere che Veblen ha offerto fondamentali contributi di riflessione a un tema oggi al centro dell'attenzione, ossia la questione dei rapporti tra sviluppo tecnologico e sviluppo sociale, fino alle implicazioni in termini di rapporto tra comunità d'affari e comunità industriale. Inoltre i suggerimenti evolutivi di Veblen appaiono non di rado d'interesse nel ricercare spiegazione a situazioni che mal si collocano all'interno della logica dell'equilibrio, come è il caso del grande fenomeno dello sviluppo territoriale ineguale. In particolare, sia sul piano generale come su quello delle specifiche applicazioni, questo saggio si sofferma a considerare recenti contributi offerti da Joan Robinson, Gunnar Myrdal e Shigeto Tsuru in tema di rilettura del pensiero di Veblen.

* Per il lettore italiano può essere interessante sapere che nel 1897 Veblen pubblicò (*Journal of Political Economy*, June 1897, vol. 5) una recensione dei *Saggi intorno alla concezione materialistica della storia* di Antonio Labriola. L'influenza dell'opera di Labriola sulla *Theory of the Leisure Class* (1899) merita probabilmente più attenzione di quanto ne abbia avuta sinora.

PROPRIETÀ DELL'EQUILIBRIO DI NASH IN MODELLI DI COMPETIZIONE TECNOLOGICA

di

FLAVIO DELBONO *

Questo lavoro intende perseguire un duplice intento. Il primo è di esaminare alcuni dei principali modelli proposti recentemente nella letteratura per analizzare situazioni in cui l'attività innovativa è oggetto di competizione tra le imprese. Tre approcci sono illustrati e discussi nei paragrafi 2, 3, 4 e 5, dopo una breve introduzione. Il secondo obiettivo è di riformulare uno dei tre approcci in esame, quello cosiddetto dinamico, al fine di renderlo più esplicitamente confrontabile con i due precedenti. Questa riesposizione del gioco dinamico – cui è dedicato il paragrafo 6 – consente di mettere a fuoco alcune interessanti proprietà dei modelli in esame.

Alcuni possibili approfondimenti sono delineati nel paragrafo conclusivo.

1. Introduzione

Nel loro influente contributo, Kamien e Schwartz (1982) identificano due approcci seguiti nella letteratura sulle relazioni tra struttura di mercato ed attività innovativa: quello di teoria delle decisioni (*decision theoretic approach*) e quello di teoria dei giochi (*game theoretic approach*).

Nel primo si considera un solo soggetto che deve scegliere, in un

* Università di Bologna, Dipartimento di Scienze Economiche e Università di Oxford, Linacre College.

Desidero ringraziare il mio Supervisor presso l'Università di Oxford, dr. John Vickers, per l'incoraggiamento e i suggerimenti dati all'elaborazione di questo lavoro. Sono grato anche ai proff. Erio Castagnoli, Piero Tani e Stefano Zamagni e ai dr. Vincenzo Denicolò e Carlo Scarpa per le osservazioni e i commenti ad una precedente stesura. Ovviamente, sono l'unico responsabile di eventuali manchevolezze ancora presenti nel lavoro.

contesto di incertezza, una strategia ottimale. Il comportamento di altri agenti – rivali effettivi o potenziali – è noto e specificato come parte delle informazioni sull'ambiente in cui il soggetto opera.

Il secondo, invece, esamina il caso in cui sono presenti almeno due agenti il cui obiettivo è di vincere una competizione (per esempio, l'ottenimento di un brevetto). Poiché questo lavoro si sofferma su alcuni modelli che si inscrivono in questo secondo approccio, conviene introdurre alcune precisazioni circa le caratteristiche di tale competizione¹.

Consideriamo un mercato di un prodotto omogeneo in cui la competizione ha per oggetto un'innovazione nel processo produttivo². Seguendo Dasgupta (1982), possiamo distinguere situazioni in cui il vincitore riceve l'intero premio (*tournaments*), da quelle in cui non si ha una netta distinzione tra vincitori e perdenti (*non-tournaments*). In quest'ultimo caso, si assume usualmente che il costo unitario di un'impresa ed il premio finale dipendano in modo continuo dal suo sforzo in R&D e talvolta anche da quello delle imprese rivali³.

La competizione nell'attività innovativa è formalizzata come un gioco non cooperativo che, sotto certe condizioni, ammette la soluzione di Nash. In una posizione di equilibrio di Nash, ogni impresa si comporta in modo ottimale, dato il comportamento delle altre imprese. Per ipotesi, tutte le imprese decidono simultaneamente la spesa in R&D o il livello di output. Poiché il gioco è simmetrico – l'output è omogeneo e i partecipanti hanno le medesime opportunità di innovare – anche l'equilibrio emergente dai modelli considerati è di tipo simmetrico (tutte le imprese adottano la stessa strategia)⁴.

Esamineremo due modelli statici ed uno dinamico. Nei primi, ogni impresa sceglie il suo livello ottimale di R&D una volta per tutte, e l'insieme delle strategie possibili è un sottoinsieme di R_+ . Nel modello esplorato nei

¹ Per una circostanziata discussione di questa modellistica, si veda KAMIEN e SCHWARTZ (1982). Si veda anche DASGUPTA (1982), REINGANUM (1984) e VICKERS (1985).

² Come accerteremo in seguito, i modelli esaminati possono essere adattati al caso di mercati con prodotti almeno potenzialmente differenziati e in cui l'innovazione riguarda il prodotto e non il processo. Si veda DIXIT e STIGLITZ (1977).

³ In quest'ultimo caso siamo in presenza dei cosiddetti effetti di *spillover* nell'acquisizione della conoscenza tecnologica. Vale a dire, la spesa in R&D di altre imprese influenza il costo unitario di produzione di una data impresa. La questione dell'interdipendenza delle funzioni di costo solleva anche il problema dell'attività imitativa. Si veda REINGANUM (1981b e 1982) per un tentativo di incorporare tale eventualità nel modello dinamico esaminato nel paragrafo 5, e SPENCE (1984).

⁴ Per una discussione ed una parziale difesa dell'ipotesi di simmetria, si rinvia a DELBONO e SCARPA (1985).

paragrafi 5 e 6, invece, ogni impresa decide l'ammontare ottimo da destinare ogni istante all'attività innovativa fino a quando si ha un'innovazione. Ogni "giocatore", dunque, non sceglie più un numero, bensì un sentiero di spesa (una funzione del tempo).

Un'altra distinzione recepita nelle prossime pagine riguarda la natura dei costi. Il primo dei due modelli statici considerati assume che i costi siano contrattuali, cioè fissi. Il secondo accoglie invece l'ipotesi che i costi siano non contrattuali, cioè che la spesa in R&D sia variabile col tempo e possa dunque essere interrotta alla data in cui si verifica un'innovazione. Il modello dinamico, da questo punto di vista, è di tipo non contrattuale.

Da ultimo, esiste incertezza tecnologica, e cioè incertezza da parte di ogni impresa circa la relazione tra la sua spesa in R&D e la data di innovazione. Comune ai modelli qui esaminati è però l'ipotesi che le aspettative delle imprese circa tale relazione non siano tra loro correlate⁵. Sulle restrizioni imposte da questa ed altre assunzioni, ci soffermeremo nel paragrafo conclusivo.

2. Il quadro di riferimento

In questo paragrafo si proverà a distinguere con precisione due formalizzazioni adottate nella letteratura per descrivere le decisioni di R&D delle imprese.

Alcuni Autori – per esempio, Reinganum (1981a, 1982) – ipotizzano che le imprese scelgano i tassi di acquisizione dell'informazione (o della conoscenza tecnologica). Sia y tale tasso per una certa impresa. Il costo di acquisizione è espresso da una funzione $c : R_+ \rightarrow R_+$, due volte differenziabile, alla quale si richiedono le seguenti – facilmente interpretabili – proprietà:

$$(i) \quad c(0) = 0$$

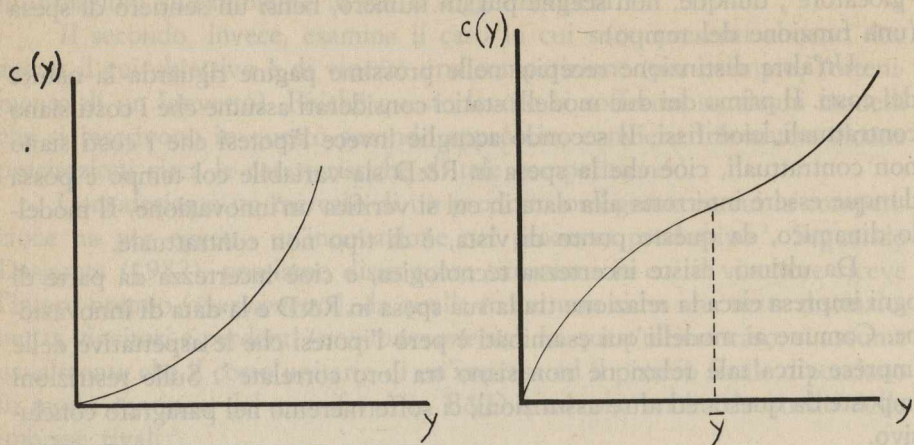
$$(ii) \quad c'(y) > 0 \text{ per } y > 0$$

$$(iii) \quad \lim_{y \rightarrow \infty} c'(y) = +\infty$$

$$(iv) \quad \exists \bar{y} \geq 0 : d^2c/dy^2 \geq 0 \quad \text{se } y \leq \bar{y}.$$

⁵ Si noti che l'incertezza tecnologica è presente anche nei modelli di "teoria delle decisioni", indipendentemente quindi dalla presenza di rivali. Si veda KAMIEN e SCHWARTZ (1982, p. 109-11).

Possibili rappresentazioni grafiche della funzione di costo sono le seguenti:



Altri Autori – per esempio, Loury (1979) e Dasgupta e Stiglitz (1980) – ipotizzano invece che le imprese scelgano i tassi di spesa in R&D.

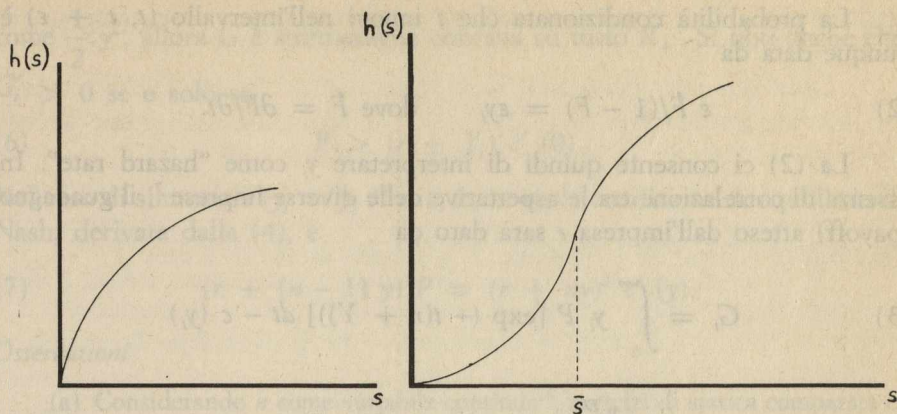
Se l'incertezza tecnologica affrontata dall'impresa è di tipo esponenziale (si veda l'espressione (1) ed i relativi commenti), possiamo esplicitare il cosiddetto "moral hazard", cioè la probabilità istantanea condizionata di un'invenzione. Tale probabilità – condizionata all'assenza di un'invenzione precedente – è indicata con $h(s)$, dove s rappresenta la spesa in R&D di una certa impresa. Ad h si richiedono le seguenti proprietà⁶:

- (i) $h(0) = 0$
- (ii) $h'(s) > 0$ per $s > 0$
- (iii) $\lim_{s \rightarrow +\infty} h'(s) = 0$
- (iv) $\exists \bar{x} \geq 0 : d^2h/ds^2 \geq 0$ se $s \geq \bar{s}$.

Possibili rappresentazioni grafiche della funzione h sono le seguenti:

⁶ h è due volte differenziabile; inoltre, trattandosi di una probabilità,

$$\lim_{s \rightarrow +\infty} h(s) = 1, \text{ dove } h : R_+ \rightarrow [0, 1].$$



Anche se le condizioni (i)-(iv) e (v)-(viii) sono visibilmente simmetriche, in quel che segue adotteremo la prima formalizzazione. Essa consente infatti di pervenire con più semplicità agli stessi risultati ottenibili con la seconda, ed inoltre ci permette di sottolineare alcune interessanti analogie tra modelli di R&D e modelli di oligopolio di tipo "quantity-setting". Giova anche precisare che nei modelli discussi nelle prossime pagine, la distinzione ora esaminata può essere eliminata semplicemente ipotizzando che y – il tasso di acquisizione dell'informazione, o sforzo in R&D – sia funzione strettamente crescente di s , la spesa in R&D.

3. Il caso contrattuale

Consideriamo $n \geq 2$ imprese che competono per introdurre un'innovazione. Il vincitore, cioè chi innova per primo, riceve un premio $P > 0$, mentre i perdenti hanno ricavi nulli. P è interpretabile come il flusso scontato di profitti valutati alla data dell'innovazione. Il tasso di sconto comune a tutti i giocatori è $r > 0$. Assumiamo inoltre che $P > c'(0)$.

Alla data $t = 0$ la generica i -esima impresa sceglie y_i – l'ammontare del suo sforzo in R&D – e sia $c(y_i)$ il costo di tale sforzo, dove c gode delle proprietà (i)-(iv) prima menzionate. Poiché stiamo considerando il caso contrattuale, $c(y_i)$ è la somma globale (lump-sum) interamente spesa alla data $t = 0$. Formalizziamo ora la presenza di incertezza tecnica. La probabilità che i introduca l'innovazione non oltre la data t è data da

$$(1) \quad F(t, y_i) = 1 - \exp(-y_i t).$$

La probabilità condizionata che i innovi nell'intervallo $(t, t + \varepsilon)$ è dunque data da

$$(2) \quad \varepsilon \dot{F}/(1 - F) = \varepsilon y_i \quad \text{dove } \dot{F} = \partial F/\partial t.$$

La (2) ci consente quindi di interpretare y_i come "hazard rate". In assenza di correlazione tra le aspettative delle diverse imprese⁷, il guadagno (payoff) atteso dall'impresa i sarà dato da

$$(3) \quad G_i = \int_0^\infty y_i P [\exp (-t(r + Y))] dt - c(y_i) \\ = \frac{y_i P}{r + Y} - c(y_i) \quad \text{dove } Y = \sum_{j=1}^n y_j.$$

Risulta agevole accertare che: la data attesa alla quale qualche impresa innova è $1/Y$; la probabilità che i innovi per primo è y_i/Y .

L'obiettivo dell'impresa i è massimizzare G_i e la sua variabile di scelta è y_i . La condizione necessaria del primo ordine (FOC) sarà

$$(4) \quad \partial G_i / \partial y_i = \frac{(r + Y_i) P}{(r + Y)} - c'(y_i) = 0$$

$$\text{dove} \quad Y_i = \sum_{j=1}^n y_j - y_i = Y - y_i$$

mentre la derivata seconda è data da

$$(5) \quad \partial^2 G_i / \partial y_i^2 = - \frac{2(r + Y_i) P}{(r + Y)^3} - c''(y_i).$$

Sappiamo dalla (iv) che G_i è strettamente concava in y_i nell'intervallo $[\bar{y}_i, \infty]$; se poi adottiamo funzioni di costo tali che $c''(y) > 0, \forall y > 0$,

⁷ In generale, la probabilità che i innovi per primo è data da $y_i^a / \sum_{j=1}^n y_j^a$ ($a \geq 1$).

Se definiamo $b = (a - 1)/a$ come indice di correlazione, assumere assenza di correlazione equivale ad ipotizzare $b = 0$, cioè $a = 1$. L'altro caso estremo, di perfetta correlazione, si ha per $a = +\infty$ che rende nullo b .

come $\frac{1}{2} y^2$, allora G è strettamente concava su tutto R_+ . Si noti anche che $G'_i > 0$ se e solo se

$$(6) \quad P > (r + Y_i) c'(0).$$

Nel caso di simmetria ($y_i = y, i = 1, 2, \dots, n$), la condizione di equilibrio di Nash, derivata dalla (4), è

$$(7) \quad (r + (n - 1) y) P = (r + ny)^2 c'(y).$$

Osservazioni

(a) Considerando n come variabile continua⁸, esercizi di statica comparata ci dicono che

$$\begin{aligned} dy/dn &< 0; \\ d(ny)/dn &> 0; \\ dy/dP &> 0. \end{aligned}$$

La prima ineguaglianza non richiede commenti. Le altre due, assieme, ci dicono che al crescere del numero di imprese, ogni impresa riduce il suo sforzo in R&D – in quanto diminuisce la probabilità individuale di “successo”, ma lo sforzo aggregato aumenta.

(b) Nel caso di perfetta libertà d'ingresso sul mercato, dobbiamo esplicitare la condizione di profitti nulli. Vale a dire, $\forall i, G_i = 0$, ovvero $Py = (r + ny) c(y)$. Quest'ultima condizione, con la (7), implica $c'(y) < c(y)/y$; in equilibrio si ha cioè eccesso di capacità in R&D⁹.

(c) Si può rilevare una stretta analogia tra questo modello di R&D ed un modello di oligopolio con fissazione delle quantità. Se y_i è reinterpretato come la quantità offerta dall'impresa i , allora la (3) altro non è che la funzione di profitto per un'impresa che offre un prodotto omogeneo in un'industria. La funzione inversa di domanda è $p(Y) = P/(r + Y)$, dove p è il prezzo ed Y è ora l'output totale dell'industria.

4. Il caso non contrattuale

Mentre nel caso contrattuale l'intera spesa in R&D è sostenuta in $t = 0$, nel caso che ora esaminiamo tale esborso si configura come un flusso di spesa che ogni impresa affronta fino al momento di un'innovazione. In

⁸ Sul significato di questa assunzione, si veda SEADE (1980, p. 482).

⁹ Su questo punto si può vedere DASGUPTA e STIGLITZ (1980), che approdano alla stessa conclusione seguendo un'altra procedura.

questo senso, la notazione precedente – si pensi a $c(y)$ – va leggermente reinterpretata¹⁰.

Come nel caso precedente, consideriamo $n \geq 2$ imprese che competono per vincere lo stesso tipo di “gara”. L’unica differenza è che il costo associato allo sforzo in R&D si compone ora di un costo fisso $C_f > 0$ e di un costo variabile (con t) che continua ad essere sostenuto fino alla data in cui qualche impresa innova. Assumendo ancora assenza di correlazione tra le aspettative incerte delle imprese, il guadagno atteso di i sarà

$$(8) \quad G_i = \int_0^{\infty} [y_i P - c(y_i)] \exp(-t(r + Y)) dt - C_f \\ = \frac{y_i P - c(y_i)}{r + Y} - C_f.$$

La FOC per determinare l’equilibrio di Nash è

$$(9) \quad \partial G_i / \partial y_i = \frac{(r + Y)(P - c'(y)) - (y_i P - c(y_i))}{(r + Y)^2} = 0$$

da cui

$$(10) \quad c'(y) = P - G_i - C_f.$$

Osservazioni

(d) $c'(0)$ ha lo stesso segno di $(P - c'(0))$, che si era assunto positivo. Quindi, $G_i \geq 0$ è condizione (necessaria) affinché il valore di y_i che massimizza G_i soddisfi la (9); diversamente, all’impresa conviene non svolgere alcuna attività di R&D e “risparmiare” almeno C_f .

(e) Se $c(y)$ è tale che G_i è strettamente concava in $y > 0$, la (9) definisce un unico massimo, a condizione che G_i sia positivo. Questa volta la condizione è necessaria e sufficiente.

(f) Se l’equilibrio (simmetrico) di Nash è stabile, esercizi di statica comparata mostrano che

$$\frac{dy}{dP} > 0; \\ \frac{dy}{dn} > 0.$$

(g) Nel caso di perfetta libertà d’ingresso, $G = 0$ ¹¹. Dalla (10) risulta che

¹⁰ Si veda LEE e WILDE (1980).

¹¹ Ciò è vero se si trascura o si esclude la presenza dei cosiddetti “sunk costs” e se $C_f = 0$.

$c'(y) > P$ e dalla (8) si ha che $yP > c(y)$. Inoltre, $c'(y) > c(y)/y$, che rovescia la relazione ottenuta nel caso contrattuale in cui si aveva "eccesso di capacità". Ovviamente, ancora, nel punto di ottimo "sociale", $c'(y) = c(y)/y$.

5. Il caso dinamico

Prima di esaminare il gioco dinamico studiato in Reinganum (1981a e 1982), risulta conveniente riesporre il modello contrattuale.

Definiamo $k_i(t)$ la conoscenza accumulata al tempo t dall'impresa i ; ricordando il significato di y_i , e assumendo $k_i(0) = 0$, si ha $k_i(t) = y_i t$. Ponendo poi $F_i(k) = 1 - e^{-k}$, la probabilità che i innovi non oltre la data t diventa $1 - \exp(-k)$, che è l'espressione (1), e il cosiddetto "moral hazard" per l'impresa i al tempo t è dato da $y_i(t)$.

Consideriamo ancora $n \geq 2$ imprese che competono per innovare. Il valore del premio è P , valutato a $t = 0$, quale che sia la data di innovazione $T < \infty$. Il tasso di sconto comune a tutte le imprese è ancora $r > 0$ ¹². Assumendo che $y_i(t)$ — il tasso di acquisizione della conoscenza — sia inferiormente limitato da $\bar{y} \geq 0$ e che $k_i(0) = 0$, avremo che $dk_i(t)/dt = y_i(t)$. Valutato al tempo t , il flusso di costo dell'acquisto dell'informazione al tasso y_i è $c(y_i)$ che particolarizziamo, per semplicità, ponendo

$$c(y) = \frac{1}{2} y^2.$$

Una strategia per l'impresa è una funzione di t . Inoltre, dato t , risulta definito anche il livello corrente di conoscenza $k(t)$. La soluzione del gioco differenziale, nel caso simmetrico, è un equilibrio di Nash in tali strategie. Le strategie di equilibrio di Nash sono del tipo:

$$(11) \quad y^*(t) = \frac{2(n-1)P \exp(rt)}{2n-1 - \exp(P(n-1)(e^{rt} - e^{rT})/r)}$$

Osservazioni

(h) Lo sforzo in R&D, y^* , è funzione crescente di t e, per un dato $t < T$, y^* è funzione crescente di n .

¹² REINGANUM (1982) assume anche che l'imitazione sia immediata e senza costo e che pure per i "perdenti" il guadagno sia positivo (non tournament). Ignoriamo queste caratteristiche al fine di rendere confrontabile il suo modello con i precedenti.

(i) Sebbene nel precedente modello t sia finito, è agevole valutare il caso di orizzonte infinito. Dalla (11) possiamo osservare che

$$\lim_{T \rightarrow +\infty} y^*(t) = \frac{2(n-1)Pe^{rt}}{2n-1};$$

vale a dire, non solo y^* cresce con t , ma tende a ∞ .

(l) Una caratteristica sorprendente del gioco differenziale ora illustrato risiede nel fatto che il guadagno derivante dall'innovazione non è scontato, mentre è scontata la spesa in R&D. Quindi, il valore dell'innovazione – valutato alla data t – è $P \exp(rt)$, e non P come nei modelli precedenti. Questa asimmetria, apparentemente ingiustificata, spiega convincentemente perché y^* cresce con t , in quanto il valore dell'innovazione aumenta esponenzialmente rispetto al costo richiesto. Ciò spiega anche perché, al limite, y^* diverge.

6. Una riformulazione del caso dinamico

In questo paragrafo si cercherà di riformulare il modello dinamico della Reinganum alla luce dell'osservazione (1) di cui sopra. Si proverà cioè a rispondere al seguente interrogativo: quali sono le implicazioni di un trattamento analogo di costi e benefici dell'attività di R&D?

Supponiamo dunque che il valore dell'innovazione per chi vince la competizione sia P , valutato alla data t , mentre manteniamo le medesime caratteristiche del modello originario. Ora la funzione (a rigore, il funzionale) obiettivo per l'impresa i sarà:

$$(13) \quad G_i = \int_0^T (y_i(t) P - \frac{1}{2} (y_i(t)^2) \exp(-rt - K(t))) dt.$$

$$\text{dove } K(t) = \sum_{j=1}^n k_j(t).$$

Il teorema di Eulero consente di scrivere

$$(14) \quad \dot{y}_i + (P - y_i)(r + Y) = y_i P - \frac{1}{2} y_i^2$$

dove $\dot{y} = \partial y_i / \partial t$.

La (14) e l'ipotesi di simmetria implicano la seguente

$$(15) \quad \dot{y} = \left(n - \frac{1}{2} \right) y^2 + (r - (n-1)P) y - rP,$$

che può essere riscritta come

$$(16) \quad 2(2n-1) \dot{y} = ((2n-1)y - \alpha - \beta)((2n-1)y - \alpha + \beta),$$

$$(17) \quad \text{dove } \alpha = (n-1)P - r$$

$$(18) \quad \beta = ((n-1)^2 P^2 + 2nrP + r^2)^{1/2}.$$

Dalla (16) si ha che

$$(19) \quad \frac{(2n-1) dy}{(2n-1)y - \alpha - \beta} - \frac{(2n-1) dy}{(2n-1)y - \alpha + \beta} = \beta dt.$$

Integrando il lato destro della (19) rispetto a y e quello sinistro rispetto a t , si ottiene

$$(20) \quad \ln \frac{(2n-1)y - \alpha - \beta}{(2n-1)y - \alpha + \beta} = \beta t + \delta,$$

dove δ è la costante di integrazione che può essere determinata dalla condizione finale

$$c'(y(T)) = y(T) = P.$$

Ciò consente di scrivere la (20) come

$$(21) \quad \ln \left[\frac{(2n-1)P - \alpha - \beta}{(2n-1)P - \alpha + \beta} \right] - \ln \left[\frac{(2n-1)y(t) - \alpha - \beta}{(2n-1)y(t) - \alpha + \beta} \right] = \beta(T - t).$$

La (21), alla luce della (17), implica la seguente

$$(22) \quad \frac{(2n-1)y(t) - \alpha - \beta}{(2n-1)y(t) - \alpha + \beta} = \frac{nP + r - \beta}{nP + r + \beta} \exp(\beta(t - T)).$$

Se indichiamo il lato destro della (22) con γ , abbiamo

$$(23) \quad y(t) = \frac{\alpha + \beta + (\beta - \alpha)\gamma}{(2n-1)(1 - \gamma)}.$$

Allorché $T \rightarrow +\infty$, ovviamente $\gamma \rightarrow +\infty$. Inoltre,

$$(24) \quad \lim_{T \rightarrow +\infty} y(t) = \frac{\alpha + \beta}{2n - 1} = \\ = \frac{P(n - 1) - r + ((n - 1)^2 P^2 + 2nPr + r^2)^{1/2}}{2n - 1}.$$

Confrontiamo ora questo risultato con quello ottenuto nel caso non contrattuale quando $c(y) = \frac{1}{2} y^2$ e $C_f = 0$. In tale caso, il guadagno atteso dall'impresa i diventa

$$(25) \quad G_i = \int_0^\infty \left(Py_i - \frac{1}{2} y_i^2 \right) \exp(-t(r + Y)) dt \\ = \frac{Py_i - \frac{1}{2} y_i^2}{r + Y}.$$

La (10) e la consueta ipotesi di simmetria implicano che

$$(26) \quad y = P - \frac{Py - \frac{1}{2} y^2}{r + ny}.$$

Inoltre,

$$(27) \quad \left(n - \frac{1}{2} \right) y^2 - ((n - 1)P - r)y - Pr = 0$$

la cui soluzione altro non è che la (24), cioè

$$(28) \quad y = \frac{\alpha + \beta}{2n - 1}.$$

Osservazioni

(m) Si è visto che se il valore dell'innovazione – così come il costo di R&D – è scontato al tasso r , la soluzione per $T \rightarrow +\infty$ è la stessa per il modello dinami-

co e per il modello non contrattuale modificato (ponendo cioè $c(y) = \frac{1}{2} y^2$ e $C_f = 0$)¹³.

(n) Risulta agevole e di qualche interesse confrontare la soluzione (28) con la soluzione di equilibrio del modello non contrattuale con orizzonte finito. L'intervallo di integrazione diventa $[0, T]$ invece che $[0, \infty]$ come nella (8) che ora diventa

$$G_i = \theta \frac{y_i P - c(y_i)}{r + Y} - C_f \quad \text{dove} \quad \theta = 1 - e^{-rT}.$$

La FOC è ancora la (17), e quando $c(y) = \frac{1}{2} y^2$ la soluzione di equilibrio è

$$\hat{y} = (\hat{\alpha} + \hat{\beta})/2n - \theta, \quad \text{dove } \hat{\alpha} = (n - \theta) P - r \\ e \hat{\beta} = ((n - \theta)^2 P^2 + 2nrP + r^2)^{1/2}.$$

La (28) è evidentemente il caso speciale in cui $\theta = 1$, cioè $T \rightarrow +\infty$.

(o) Qual è il significato dell'equivalenza tra il modello non contrattuale e quello dinamico? Se anche il premio è opportunamente scontato e l'orizzonte temporale è infinito, il gioco è sostanzialmente stazionario. Se nessuna impresa riesce ad innovare nell'intervallo $[0, \varepsilon)$, allora il sotto-gioco che inizia alla data $t = 0$ non è altro che una copia del gioco iniziato alla data $t = 0$. Risulta allora naturale attendersi che le strategie di equilibrio siano le stesse per gli intervalli $[0, \varepsilon)$ e $[\varepsilon, 2\varepsilon)$, e se questo si verifica per ogni ε positivo, allora le strategie sono costanti.

7. Osservazioni conclusive

In queste considerazioni finali non si proverà a duplicare le osservazioni proposte alla fine dei paragrafi precedenti. Piuttosto, si cercherà di delineare brevemente alcune possibili estensioni di questi approcci.

Nell'illustrare i tre modelli considerati, si è talvolta parlato impropriamente

¹³ Si noti però la complessità del modello dinamico, che rende particolarmente complicata la derivazione della soluzione anche nel caso speciale in cui $c(y) = \frac{1}{2} y^2$ o, più in generale, per funzioni di costo isoelastiche, del tipo

$$c(y) = Ay^{-B} \quad \text{con } A, B > 0.$$

Per funzioni di costo siffatte, B è interpretabile come un indice dell'opportunità innovativa; si veda DASGUPTA (1982, p. 22). Una dimostrazione alternativa dell'equivalenza tra i due modelli, nel caso di orizzonte temporale infinito, è illustrata in Appendice.

mente di *gara* tra le imprese. Impropriamente, poiché una rigorosa e più soddisfacente analisi di gare tra due o più partecipanti dovrebbe incorporare almeno due aspetti qui trascurati.

Il primo riguarda la "memoria" circa il passato del gioco, che una distribuzione esponenziale non consente di incorporare. Vale a dire, l'esperienza acquisita dalle imprese nella storia passata non gioca alcun ruolo nelle decisioni presenti¹⁴. In questi modelli, infatti, il cosiddetto "hazard rate" dipende solo dai livelli correnti di sforzo in R&D.

Il secondo aspetto di insoddisfazione concerne l'ipotesi di assenza di correlazione tra le aspettative delle imprese. Una più realistica descrizione di una gara — cioè di una situazione in cui si scontrano diversi comportamenti strategici — dovrebbe consentire almeno una parziale correlazione tra le aspettative delle imprese rivali. Un'investigazione dei presenti modelli secondo queste linee ha formato l'oggetto di ulteriore ricerca e i primi risultati sembrano incoraggianti¹⁵.

Altri approfondimenti sembrano meritare ulteriori indagini. Un primo motivo di interesse muove dall'osservazione che la attività di R&D è soltanto una delle strategie accessibili all'impresa, così come la concorrenza tecnologica è solo una delle possibili forme di competizione. Le integrazioni tra le decisioni di R&D e le altre scelte strategiche dell'impresa costituiscono un capitolo importante dei nuovi approcci ai problemi della cosiddetta economia industriale.

Questi temi — così come l'esame della relazione tra struttura industriale e concorrenza tecnologica — sembrano però richiedere un'analisi a livello di industria, più che di singola impresa¹⁶.

Un'ulteriore caratteristica della modellistica presa in esame è l'assenza degli aspetti finanziari dell'attività innovativa¹⁷. La capacità dell'impresa di svolgere tale attività può però dipendere talvolta più dalla sua situazione finanziaria che dal grado di efficienza della sua produzione di R&D. Anche

¹⁴ Questo aspetto del gioco riguarda apparentemente la distinzione tra giochi a completa o incompleta informazione, per la quale si rinvia, ad esempio, ad HARSANYI (1977). Apparentemente, poiché di fatto, nei primi due modelli qui considerati, le decisioni passate non influenzano in alcun modo la probabilità "attuale" di successo. Il modello della Reinganum, invece, assume che la probabilità istantanea di introdurre un'innovazione sia funzione sia del tasso corrente di investimento, sia dello stock accumulato di conoscenza tecnica.

¹⁵ A questo proposito, mi permetto di rinviare ancora a DELBONO e SCARPA (1985), che contiene anche alcuni rilievi critici sulle limitazioni più vistose di questa modellistica. Si veda anche SCARPA (1985).

¹⁶ Per un interessante tentativo in questa direzione, si veda DASGUPTA e STIGLITZ (1981).

¹⁷ Sulla questione si veda ROBERTS e WEITZMAN (1981); sulle barriere finanziarie si può vedere ZILIOITI (1984).

questo ultimo punto, a parere di chi scrive, dovrebbe formare oggetto di ulteriore riflessione.

APPENDICE ¹⁸

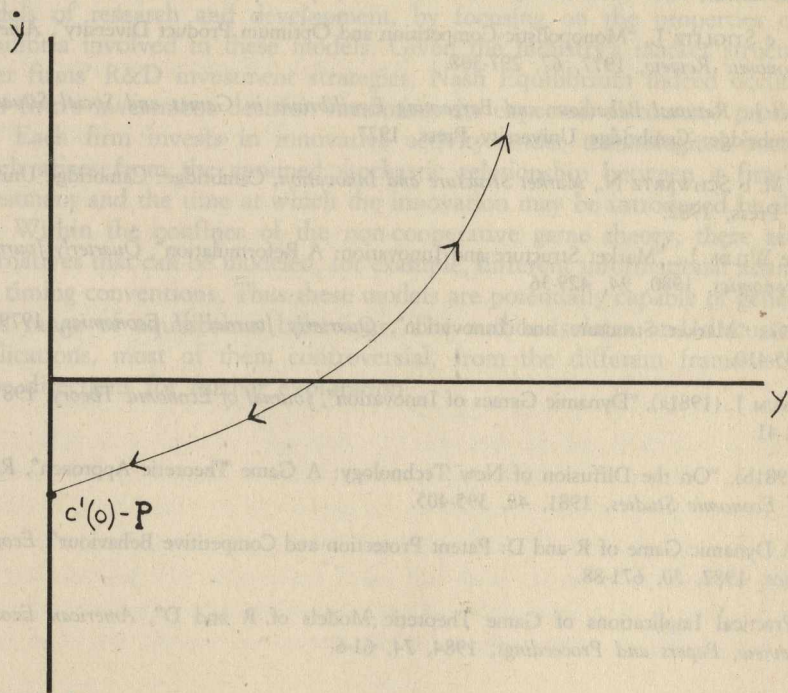
Nel caso di orizzonte temporale infinito, se la funzione di costo è strettamente convessa, possiamo generalizzare la dimostrazione – presentata nel paragrafo 6 – dell'equivalenza tra il caso non contrattuale e il caso dinamico. La (13) può essere infatti riscritta come

$$(29) \quad G_i = \int_0^{\infty} [y_i P - c(y_i)] \exp(-rt - k) dt.$$

L'equazione di Eulero consente di scrivere

$$(30) \quad c''(y_i) \dot{y}_i = [Py_i - c(y_i)] - (r + Y)(P - c'(y_i)).$$

Assumendo che il punto di equilibrio simmetrico (cioè, $Y = ny_i$) sia unico ¹⁹, dato che $c'' > 0$, il seguente diagramma di fase mostra che l'equilibrio è instabile.



¹⁸ La dimostrazione contenuta in questa Appendice mi è stata suggerita da V. Denicolò.

¹⁹ Si assume cioè che esista un unico valore di y_i per cui $\dot{y}_i = 0$; si veda anche la precedente osservazione (e).

In altri termini, con un orizzonte temporale infinito, la soluzione del modello dinamico deve coincidere con il punto di equilibrio statico²⁰. Infatti, ponendo $\dot{y}_i = 0$ nella (30), si ottiene

$$(31) \quad Py_i - c(y_i) = (r + Y)(P - c'(y_i))$$

che è proprio la condizione che si evince dalla (9).

RIFERIMENTI BIBLIOGRAFICI

- DASGUPTA P., "The Theory of Technological Competition", *ICERD Discussion Paper*, London School of Economics, 1982.
- e STIGLITZ J., "Uncertainty, Industrial Structure and the Speed of R&D", *Bell Journal of Economics*, 1980, 11, 1-28.
- and —, "Entry, Innovation, Exit", *European Economic Review*, 1981, 23, 137-58.
- DELBONO F. e SCARPA C., "Aspettative, competizione tecnologica e ottimalità dell'attività innovativa", *Note Economiche*, 3-4/1985, 78-103.
- DIXIT A. e STIGLITZ J., "Monopolistic Competition and Optimum Product Diversity", *American Economic Review*, 1977, 67, 297-308.
- HARSANYI J., *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge: Cambridge University Press, 1977.
- KAMIEN M. e SCHWARTZ N., *Market Structure and Innovation*, Cambridge: Cambridge University Press, 1982.
- LEE T. e WILDE L., "Market Structure and Innovation: A Reformulation", *Quarterly Journal of Economics*, 1980, 94, 429-36.
- LOURY G., "Market Structure and Innovation", *Quarterly Journal of Economics*, 1979, 93, 395-410.
- REINGANUM J. (1981a), "Dynamic Games of Innovation", *Journal of Economic Theory*, 1981, 25, 21-41.
- (1981b), "On the Diffusion of New Technology: A Game Theoretic Approach", *Review of Economic Studies*, 1981, 48, 395-405.
- , "A Dynamic Game of R and D: Patent Protection and Competitive Behaviour", *Econometrica*, 1982, 50, 671-88.
- , "Practical Implications of Game Theoretic Models of R and D", *American Economic Review, Papers and Proceedings*, 1984, 74, 61-6.

²⁰ Il ragionamento è analogo a quello usualmente adottato per dimostrare che si deve scegliere la condizione iniziale in modo tale da situarsi sul ramo stabile di un punto di sella. Il nostro equilibrio instabile è infatti un punto di sella "degenere".

- ROBERTS K. e WEITZMAN M.L., "Funding Criteria for Research, Development and Exploration Projects", *Econometrica*, 1981, 49, 1261-88.
- SCARPA C., "Innovazioni, esternalità e barriere alla entrata in mercati altamente innovativi", *Ricerche Economiche*, 1985, 39, 1-21.
- SEADE J., "On the Effects of Entry", *Econometrica*, 1980, 48, 479-89.
- SPENCE M., "Cost Reduction, Competition and Industry Performance", *Econometrica*, 1984, 52, 101-21.
- VICKERS J., "Patent Races and Market Structure", Tesi di D. Phil. non pubblicata, 1985, Oxford.
- ZILIOTTI M., "Le barriere finanziarie all'entrata", *Quaderni dell'Istituto di Scienze Economiche*, Università di Parma, 1984.

PROPERTIES OF NASH EQUILIBRIUM IN R&D MODELS

The main purpose of this paper is to investigate some recent game theoretic models of research and development, by focusing on the properties of Nash Equilibria involved in these models. Given the industry's market structure and other firms' R&D investment strategies, Nash Equilibrium indeed occurs when each firm's investment decision maximizes its expected discounted profits.

Each firm invests in innovative activity under technological uncertainty, which arises from the assumed stochastic relationship between a firm's R&D investment and the time at which the innovation may be introduced by the firm.

Within the confines of the non-cooperative game theory, there are many alternatives that can be modeled, for example, different informational assumptions and timing conventions. Thus these models are potentially capable of generating a wide range of equilibrium behaviour. The author isolates and discusses some implications, most of them controversial, from the different frameworks, and some directions for further exploration.

CAUSES AND CONSEQUENCES OF LABOR MILITANCY

or

THE COMMON FACTOR IN INDUSTRIAL CONFLICT AND UNIONIZATION DYNAMICS

by

MANFRED GÄRTNER * and ERWIN W. HERI *

Abstract

Investigations into the interaction between industrial conflicts, trade union membership levels or rates of change on the one hand and wage dynamics on the other have implicitly assumed or explicitly postulated that one of the first three variables may be used as a measuring rod for some rather vague, unobservable concept named labor militancy. Numerous articles have attempted to establish the superiority of one of these competing indicators over the others. In order to avoid having to make a rather arbitrary choice, this paper employs Jöreskog's LISREL framework to make use of all three indicators of labor militancy simultaneously.

Pathbreaking insights into the economic aspects of collective militancy or conflict are due to Zeuthen (1930) and Hicks (1932). While these authors remain vague on the implied direction of causality, subsequent empirical work does rather explicitly postulate certain causal patterns. Empirical analyses conducted in the context of the Phillips curve proceed from the assumption that organized labor militancy is an autonomous force behind the wage determination process. In other words, it is maintained that wage and (by means of some sort of mark-up-pricing assumption) price

* University of Basel (Switzerland).

We are particularly grateful to George M. Sheldon and Bruce B. Fitzgerald for their constructive criticisms and helpful suggestions.

inflation are a *consequence* of labor militancy¹. A second and typically separately thriving branch of analysis considers real wage inflation a major *cause* of manifest labor militancy². It is one purpose of this paper to formulate a fairly general model describing the dynamic *interaction* between labor militancy and real wage movements.

The second purpose deals with the fact that empirical investigations into the causes and consequences of labor militancy have always been based on a more or less arbitrary and highly disputed choice from among a variety of available indicators and proxies. To name but a few of many pertinent studies, Dicks-Mireaux and Dow (1959, pp. 122, 145 ff.) have augmented their Phillips-type wage equation by their own *subjective index* of labor militancy. Hines (1964) produced some rather surprising statistical results by postulating that the *change in the percentage of the labor force unionized* is a prime indicator for trade union militancy. Many others did and do consider *strike activity* the most reasonable and reliable measure of militancy of organized labor.

Each of these measures is disputed and has its deficiencies³. But each has its merits, as well, and the selection of one and the neglect of some other may deprive an empirical model of valuable information on the militancy variable of the theoretical model. The approach adopted in the empirical section of the present paper, therefore, is to treat labor militancy as a latent, unobservable concept, and to use unionization and strike series as indicators. The model is successfully estimated with West German data for the time period 1952 to 1981.

1. *A model of the causes and consequences of labor militancy in organized labor markets*

The model to be proposed is basically an extension and a generalization of the seminal Ashenfelter and Johnson (1969) model, which has dominated analyses of strike activity ever since its publication more than a decade ago⁴. Collective wage bargaining is considered as an interplay between three groups or actors: union leaders, employers' representatives, and the union rank and file. While the first two conduct the actual wage bargain-

¹ Studies analyzing the role of militancy in the wage determination process are surveyed in GÄRTNER (1981, 1983).

² For references to and results of this approach see GÄRTNER (1981, 1983).

³ See PURDY and ZIS (1974), STERN (1978) and EDWARDS (1979).

⁴ The model also draws on the first author's recent strike model (GÄRTNER, 1985).

ing, the union rank and file is only indirectly involved by restricting the behavior of their leadership. This restriction operates as follows: union leaders are assumed to be utility maximizers in the sense that they try to ensure their survival as elected representatives of their trade union and also the survival and further growth of their trade union⁵. Both these objectives are enhanced by a performance at the bargaining table which does not disappoint rank and file aspirations. If, therefore, at the point of expiration of the previous contract, rank and file aspirations exceed the wage increase offered by the management, union leaders will choose to resort to some sort of revealed labor militancy in order to moderate rank and file aspirations by driving home the plain fact that initially claimed wage hikes cannot be met.

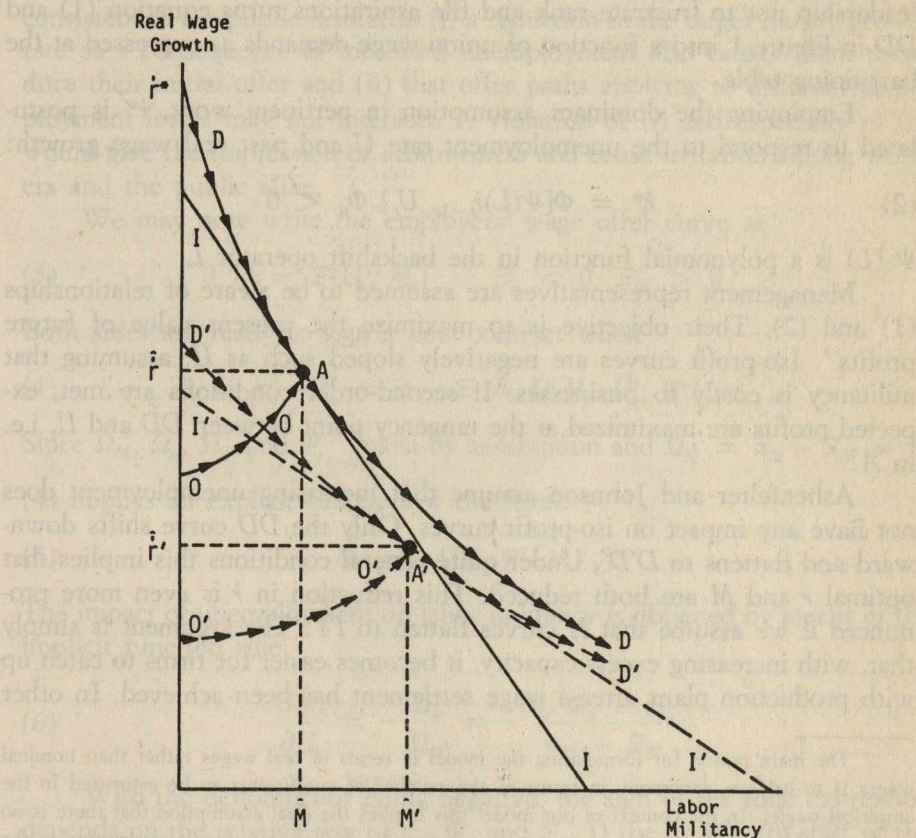


FIGURE 1: Illustration of Labor Militancy Model

⁵ This general view of union leadership is due to Ross (1948). It also serves as a central building block in ASHENFELTER and JOHNSON (1969).

Escalation of militancy is purposely continued until the gap between the management's offer and the rank and file claim has been eliminated.

A graphical illustration of this mechanism is given in Figure 1. The *DD*-curve depicts the moderating influence of escalating labor militancy on rank and file aspirations. In formal terms

$$(1) \quad \dot{r}_t^D = \pi(\dot{r}_t^*, M_t) \pi_{\dot{r}^*} > 0; \pi_M < 0; \pi(\dot{r}_t^*, 0) = \dot{r}_t^*$$

\dot{r}^D denotes the rate of change of real wages expected by rank and file members which, according to the above argument, is also the wage growth rate *demanded* by union leadership⁶. \dot{r}^* denotes rank and file aspirations held at the beginning round. M is labor militancy. The strategy of union leadership not to frustrate rank and file aspirations turns equation (1) and *DD* in Figure 1 into a function of union wage demands as expressed at the bargaining table.

Employing the dominant assumption in pertinent work, \dot{r}^* is postulated to respond to the unemployment rate U and past real wage growth:

$$(2) \quad \dot{r}_t^* = \Phi[\Psi(L)\dot{r}_{t-1}, U_t] \quad \Phi_u < 0$$

$\Psi(L)$ is a polynomial function in the backshift operator L .

Management representatives are assumed to be aware of relationships (1) and (2). Their objective is to maximize the present value of future profits⁷. Iso-profit curves are negatively sloped such as *II*, assuming that militancy is costly to businesses. If second-order conditions are met, expected profits are maximized at the tangency point between *DD* and *II*, i.e. in *A*.

Ashenfelter and Johnson assume that increasing unemployment does not have any impact on iso-profit curves. Only the *DD* curve shifts downward and flattens to *D'D'*. Under quite general conditions this implies that optimal \dot{r} and M are both reduced. This reduction in \dot{r} is even more pronounced if we assume that *II* curves flatten to *I'I'*. The argument is simply that, with increasing excess capacity, it becomes easier for firms to catch up with production plans after a wage settlement has been achieved. In other

⁶ The main reason for formulating the model in terms of real wages rather than nominal wages is to achieve parsimony in terms of the number of coefficients to be estimated in the empirical model. In the context of our model this implies the dual assumption that there is no money illusion and that inflation expectations are rational. Tests of these assumptions in the framework of a strike study and with a similar sample period as the one to be used below, led to very encouraging results (see GÄRTNER, 1985).

⁷ For a more detailed description of this and a formal derivation of conditions which maximize the present value of future profits, see ASHENFELTER and JOHNSON (1969).

words, a one percentage point "saving" in real wage growth must be matched by a higher level of costly militancy.

While this determines the foreseen and optimal bargaining *outcome* to be \dot{r} after militancy has increased to \hat{M} , it does not state how employers' offers *develop* as M increases. It is clear that no offer lying *above* II will be made. In reality employers appear to begin with an initial wage offer *below* \hat{r} . This allows for the possibility to "give in" to union demands as militancy escalates and finally agree to expected profit maximizing wage growth \hat{r} as an "honest" compromise. An arbitrary wage offer curve OO based on this argument is given in Figure 1.

The actual shape of the offer path OO has no impact on contracted real wage growth⁸. We, hence, merely assume that employers attempt to appear consistent to the public insofar as (i) a reduction in the target money growth rate as a consequence of increased unemployment also causes them to reduce their initial offer and (ii) that offer paths applying to different unemployment levels may not intersect. A violation of (i) and especially of (ii) would give the impression of arbitrariness and cause irritation among workers and the public alike.

We may now write the employers' wage offer curve as

$$(3) \quad \dot{r}_t^o = \Xi(U_t, M_t) \quad \Xi_u < 0, \Xi_M > 0$$

Both sides are ready to sign a new contract when

$$(4) \quad 0 = \dot{r}_t^D - \dot{r}_t^o = \pi - \Xi \equiv \Omega[M_t, U_t, \Psi(L)\dot{r}_{t-1}]$$

Since Ω_M , Ω_u , Ω_Ψ and $\Psi_{r_{t-i}}$ exist by assumption and $\Omega_M = \pi_M - \Xi_M \neq 0$,

(4) implies an explicit function of the form

$$(5) \quad M_t = \omega[U_t, \Psi(L)\dot{r}_{t-1}].$$

The impact of unemployment on labor militancy is obtained by means of the implicit function rule:

$$(6) \quad \frac{\partial M_t}{\partial U_t} = - \frac{\Omega_u}{\Omega_M} = - \frac{\pi_r \Phi_u - \Xi_u}{\pi_M - \Xi_M}$$

With the denominator being negative, the sign of the total expression depends on the relative size of π_r , Φ_u and Ξ_u . If the downward-shift of the \dot{r}^D schedule exceeds the downward shift of the \dot{r}^o schedule, the militancy level drops. If the opposite relation holds, militancy escalates.

* Except for the case when management holds an erroneous expectation of the DD locus.

The Ashenfelter and Johnson-model, which may be conceived as a special case of the above model, operates with the assumption $\Xi_u = 0$. This results in a negative relationship between M and U . The literature on strikes has literally internalized this trade-off notion without paying any attention to this particular assumption on which it is based. At the opposite extreme we might assume $\Phi_u = 0$, which produces a quite different hypothesis: increasing militancy with increasing unemployment.

What we need in this situation are conditions under which Ξ_u or Φ_u may be expected to be small. We postulate that the magnitude of Φ_u highly depends on the financial consequences of becoming unemployed. This implies the assumption that in welfare states the probability of Φ_u being small and of $\pi^* \Phi_u$ falling short of Ξ_u in magnitude is high, which means that $\partial M_t / \partial U_t > 0$. The opposite has to be expected for countries with a comparatively less developed public welfare system⁹.

The response of labor militancy to recent real wage growth is given by

$$(7) \quad \frac{\partial M_t}{\partial \dot{r}_{t-i}} = - \frac{\Omega \dot{r}_{t-i}}{\Omega_M} = - \frac{\pi_{r^*} \Phi_\Psi \Psi r_{t-i}}{\pi_M - \Xi_M}.$$

Again the sign of $\partial M_t / \partial \dot{r}_{t-i}$ is indeterminate, given the assumptions we have made so far. If $\Psi r_{t-i} > 0$ we obtain a positive relationship, if $\Psi r_{t-i} < 0$, a negative one results.

Strike analyses unanimously postulate $\partial M_t / \partial \dot{r}_{t-i} < 0$. This seems to rest on a rationalization given by Ashenfelter and Johnson which has the peculiar property that the acceptable money wage increase at the point of contract expiration is zero if the currently anticipated increase in real wages equals the expected long-run increase. Assume here that the acceptable wage increase regresses to expected long-run real wage growth \dot{r}_t^L . A fairly general formulation of this is

$$(8) \quad \dot{r}_t^* = \dot{r}_t^L - \sum_{i=1}^n \gamma_i (\dot{r}_{t-i} - \dot{r}_t^L)$$

Assume for simplicity that long-run expectations formation is based on the same maximum lag length n as the regressive memory, i.e.

$$(9) \quad \dot{r}_t^L = \Psi^n(L) r_{t-1}$$

After substitution of (9) into (8) the coefficient of r_{t-i} turns out to be

⁹ We would assume countries such as Canada, the United Kingdom, or the United States to belong to this second category. Typical for the first category seem to be some Scandinavian welfare states, the Netherlands, or West Germany, the latter being the object of our empirical analysis.

$\Psi_i + \Psi_i \gamma_i - \gamma_i$. While the sign of each individual lag coefficient will obviously depend on the particular lag distributions of Ψ and γ , the fact that both $\Sigma\Psi$ and $\Sigma\gamma$ should not be too far from 1 gives a strong positive bias to expected coefficient signs. E.g., if last period's real wage growth is expected to continue in the long run ($\Psi_1 = 1$) and regression of actual to long-run wage growth is assumed to be full and completed within one period ($\gamma_1 = 1$), then $\Psi_i + \Psi_i \gamma_i - \gamma_i = 1$ for $i = 1$ and zero for all other i . The same qualitative result obtains if $n > 1$, unless we are dealing with rather implausible lag distributions. We therefore conclude that a set of fairly reasonable and general assumptions produces a *positive* impact of past real wage growth on present labor militancy.

2. A latent variable approach to labor militancy in West Germany

This section presents the results of an empirical analysis of West German labor militancy. The analysis draws on a linearized version of the above model. In order to avoid the customarily arbitrary choice of some single proxy of militancy, it is considered here as a latent, inherently unobservable concept. An appropriate estimation technique is based on an analysis of the covariance structure of the model presented in section 1¹⁰.

We shall now briefly describe a general method of covariance analysis for estimating the coefficients of a set of Linear Structural Relations (LISREL)¹¹, the application of this general approach to our labor militancy model and its estimation by maximum likelihood methods.

2.1. The general LISREL approach

The variables in the equations system of a LISREL model are latent (unobservable) variables, and there are multiple indicators, causes and consequences of each latent variable. The approach allows for errors in equations (residuals, disturbances) as well as for errors in the observed variables (errors of measurement) and yields estimates of the residual covariance matrix and the measurement covariance matrix¹². It further produces

¹⁰ Analysis of covariance structures has become a common term for a number of techniques for analyzing multivariate data in order to detect and assess unobservable sources of variation and covariation in observed indicators. See JÖRESKOG (1981).

¹¹ See JÖRESKOG and SÖRBOM (1981).

¹² In many applications the origin and the unit in the scale of measurement are arbitrary or irrelevant and then only the correlation matrix may be of interest. This is the approach we used in our work. This does, however, not alter the general assertions, as LISREL works equally with covariance, correlation or moment matrices (see JÖRESKOG and SÖRBOM, 1981).

an estimate of the matrix of structural coefficients. The method covers a wide range of models useful in the social sciences. Although it was originally designed to handle models with latent variables, measurement errors and reciprocal causation, it can also be used in a traditional econometric framework, placing special assumptions on the error matrices. This point will be clarified by the following formal description of the statistical model.

The LISREL model considers random vectors $\eta' = (\eta_1, \eta_2, \dots, \eta_m)$ and $\xi' = (\xi_1, \xi_2, \dots, \xi_n)$ of latent dependent and independent variables, respectively. These are related by a system of linear structural relations,

$$(10) \quad B\eta = \Gamma\xi + \zeta$$

where B ($m \times m$) and Γ ($m \times n$) are coefficient matrices and ζ ($m \times 1$) is a vector of random disturbance terms. It is assumed that $E[\zeta] = 0$ and that ξ is uncorrelated with ζ and, furthermore, that B is nonsingular.

The vectors η and ξ are *unobservable*. Instead we can *observe* vectors $y' = (y_1, y_2, \dots, y_p)$ and $x' = (x_1, x_2, \dots, x_q)$ such that

$$(11) \quad \begin{aligned} y &= \Lambda_y \eta + \varepsilon \\ x &= \Lambda_x \xi + \delta \end{aligned}$$

where ε and δ are vectors of errors arising from the measurement of y and x , respectively. We assume further that y and x are measured as deviations from their means, which implies $E(y) = E(x) = 0$. The matrices Λ_y ($p \times m$) and Λ_x ($q \times n$) contain regression coefficients of y on η and of x on ξ , respectively. Measurement errors are assumed to be uncorrelated with the latent variables η and ξ and the disturbance term ζ , but may be correlated among themselves.

Let Φ ($n \times n$), Ψ ($m \times m$), θ_ε ($p \times p$) and θ_δ ($q \times q$) be the covariance matrices of ξ , ζ , ε and δ , respectively. From the above assumptions it follows, that the covariance matrix $\Sigma[(p + q) \times (p + q)]$ of $z = (y', x')$ is

$$(12) \quad \Sigma = \begin{bmatrix} \Lambda_y B^{-1} (\Gamma \Phi \Gamma' + \Psi) B^{-1} \Lambda_y + \theta_\varepsilon & \Lambda_y B^{-1} \Gamma \Phi \Lambda_x' \\ \Lambda_x \Phi \Gamma' B^{-1} \Lambda_y' & \Lambda_x \Phi \Lambda_x' + \theta_\delta \end{bmatrix}$$

All the elements of Σ are functions of the elements of the eight matrices Λ_x , Λ_y , B , Γ , Φ , Ψ , θ_ε , θ_δ . In applied work a number of these elements will – for reasons of identification – have to be fixed and set equal to predetermined values.

Equations (10) and (11) define – together with the accompanying assumptions – the general LISREL model.

Let z_1, z_2, \dots, z_m , be independent observations of $z = (x', y')$ and let $\bar{z} = (\bar{y}, \bar{x})$ be the vector of sample means, and

$$(13) \quad S = \frac{1}{N} \sum_{a=1}^N (z_a - \bar{z}) (\bar{z}_a - \bar{z})'$$

the sample covariance matrix (with N = Number of observations minus 1).

Furthermore let θ be the vector of all free parameters in all the eight parameter matrices above. The LISREL approach (Jöreskog and Sörbom, 1981) is to fit the model to the input data for x and y by minimizing the function

$$(14) \quad F[\Sigma(\theta)] = \log |\Sigma| + \text{tr}(\Sigma^{-1} S) - \log |S| - (p + q)$$

with respect to θ . If the distribution of the observed variables is multivariate-normal, this yields maximum likelihood estimates which are efficient in large samples.

The minimum value of F times the sample size applies to be a measure of goodness-of-fit of the whole model. This measure is distributed as χ^2 with $1/2 (p + q) (p + q + 1) - t$ degrees of freedom, where t stands for the order of θ . The standard errors of the estimates can be obtained from the matrix of expected second order derivatives of F (information matrix).

2.2. The model of labor militancy in LISREL terminology

The model presented in the first section of this paper may account for

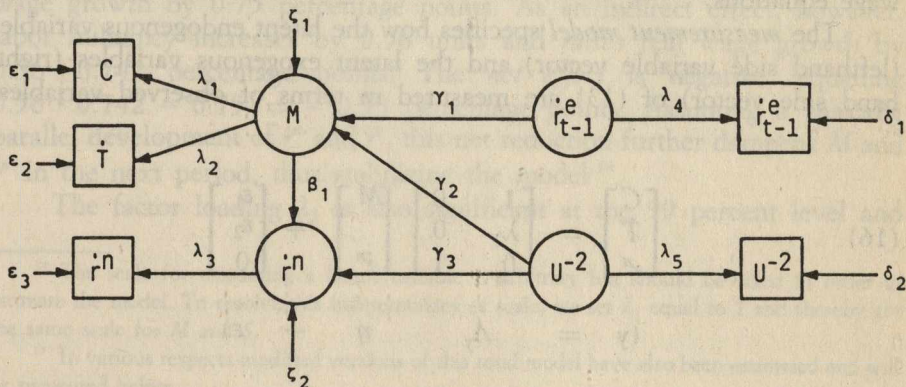


FIGURE 2: Path diagram of the main model of the empirical analysis of the causes and consequences of labor militancy (Version 1); general form, restrictions not included.

both labor militancy [equation (5)] and real wage growth [equation (3)]. Linearizing both equations and employing some additional assumptions to be mentioned below leads to the path diagram in Figure 2. The diagram emphasizes that our LISREL model consists of two submodels: (i) the *structural equation model* which relates the endogenous and exogenous variables of the theoretical model to each other [as in (10)]. These variables are circled to indicate that they are considered latent. (ii) The *measurement model*, which relates all latent concepts of the structural equation model to one or more observable indicators [as in (11)]. These are boxed in Figure 2.

In matrix notation the obtained *structural equations model* reads

$$(15) \quad \begin{bmatrix} 1 & 0 \\ -\beta_1 & 1 \end{bmatrix} \begin{bmatrix} M \\ \dot{r}'' \end{bmatrix} = \begin{bmatrix} \gamma_1 & \gamma_2 \\ 0 & \gamma_3 \end{bmatrix} \begin{bmatrix} \dot{r}_{t-1}'' \\ U^{-2} \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ z_2 \end{bmatrix}$$

$$(B \quad \eta = \Gamma \quad \xi + \zeta)$$

Note that in empirical applications our model calls for the use of different wage series in both equations: while equation (3) explains the bargaining outcome (i.e. *negotiated* wage growth \dot{r}''), union rank and file aspirations as used in (5) respond to past movements of effective real wages (\dot{r}''). It is also assumed that only real wage growth lagged by *one* year is relevant. This assumption has proven to be reasonable and successful in empirical analyses of West German strike data¹³. It also has the positive side effect of making the model parsimonious in terms of coefficients to be estimated. Finally, the unemployment rate is given a power of -2 in order to produce non-linear wage growth. This fits well with worldwide experience with Phillips-type wage equations.

The *measurement model* specifies how the latent endogenous variables (lefthand side variable vector) and the latent exogenous variables (right-hand side vector) of (13) are measured in terms of observed variables:

$$(16) \quad \begin{bmatrix} C \\ \dot{r}'' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} M \\ \dot{r}'' \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ 0 \end{bmatrix}$$

$$(y = \Lambda_y \quad \eta + \varepsilon)$$

¹³ Nevertheless, we have also tested a variant of the model which includes \dot{r}_{t-2}'' as an additional explanatory variable in the structural equations model. The result is discussed below.

$$(17) \quad \begin{bmatrix} \dot{r}_{t-1}^e \\ U \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{r}_{t-1}^e \\ U \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(X = \Lambda_x \quad \xi + \delta)$$

The focus of the present study is on the role of labor militancy as a latent concept. Other variables are considered directly observable and free of measurement errors. This is taken care of by appropriate restrictions on Λ_y , Λ_x , ε and δ ¹⁴. Indicators for latent militancy are industrial conflict C and the change in union membership \dot{T} .

Equations (16) to (17) and the assumption that Θ_ε and Ψ are diagonal constitute a simple and – as will be demonstrated in the following section – rather successful variant of our model of the causes and consequences of labor militancy¹⁵.

2.3. Estimation results

Estimates of the parameters of our model are obtained by fitting the covariance matrix Σ implied by the *model* – as given in appendix A – to the *sample* covariance matrix S calculated from our data. The sample period is 1952 to 1981, annual observations. Definitions and sources of data series are given in appendix B.

Estimation results are summarized in Table 1. Reference to Figure 2 or equations (16) to (17) will facilitate the economic interpretation of the results. All coefficients of the structural model are significant at the 99 percent level. Rising unemployment appears to have a dual effect on real wage growth. As a *direct* effect, a reduction of U^{-2} by one unit *reduces* real wage growth by 0.75 percentage points. As an indirect effect, however, labor militancy increases by 0.56 units and *raises* real wage growth by $0.56 \cdot 0.742$ percentage points. The *net effect* is negative, equalling $0.56 \cdot 0.742 - 0.75 = -0.33$ percentage points. Assuming a basically parallel development of \dot{r}^m and \dot{r}^e , this net reduction further dampens M and \dot{r}^m in the next period, thus stabilizing the model¹⁶.

The factor loading λ_2 is also significant at the 99 percent level and

¹⁴ The scale for measuring a latent variable is arbitrary but should be fixed in order to estimate the model. To resolve this indeterminacy of scale, we set λ_1 equal to 1 and thereby got the same scale for M and S .

¹⁵ In various respects modified versions of this total model have also been estimated and will be presented below.

¹⁶ Real wage growth rates follow the first-order difference equation $(1 - 0.460L) r_t = 0.334 U_t^{-2}$, which converges monotonically.

TABLE 1.

CAUSES AND CONSEQUENCES OF LABOR MILITANCY
IN WEST GERMANY; VERSION 1; ANNUAL OBSERVATIONS 1952-1981;
LISREL ESTIMATES

		Estimations	Standard Errors
Coefficients of structural equations and measurement models	β_1	0.742 **	0.264
	γ_1	0.620 **	0.167
	γ_2	- 0.560 **	0.163
	γ_3	0.750 **	0.154
	λ_1	1.000	c
	λ_2	0.966 **	0.306
	$\lambda_3, \lambda_4, \lambda_5$	1.000	c
Error variances	$V[\varepsilon_1]$	0.500 **	0.172
	$V[\varepsilon_2]$	0.534 **	0.176
	ψ_1	0.158	0.116
	ψ_2	0.433 **	0.133
Goodness of fit statistics	χ^2 (d.f.)	1.20 (3)	
	p	0.75	
	AGFI	0.93	
	CDY	0.61	
	CDS	0.85	

Notes:

* Indicates significance at the 95 percent level, two-tailed test.

** Indicates significance at the 99 percent level, two-tailed test.

c The parameter has been constrained to the given value.

χ^2 Likelihood ratio test statistic for testing the hypothesis that Σ is of the form implied by the model [for caveats in using the chi-square test see Jöreskog and Sörbom (1981) and Bentler (1982)].

p Probability of getting a chi-square value exceeding the one actually obtained, given that the postulated constraints are true.

AGFI Adjusted goodness-of-fit index, computed as

$$AGFI = 1 - \left[\frac{\text{tr}(\Sigma^{-1} S - I)^2}{\text{tr}(\Sigma^{-1} S)^2} \frac{K(K+1)}{2d} \right]$$

K Denotes the number of observed variables and d are degrees of freedom of the model [see Jöreskog and Sörbom (1981)].

CDY $\equiv 1 - |\theta|/|S|$ is the coefficient of determination of the y -variables.

CDS $\equiv 1 - |\psi|/|\text{Cov}(\eta)|$ is the coefficient of determination of the structural equations model.

p , AGFI, CDY and CDS are restricted to the interval between 0 and 1 and increase as a model improves.

indicates that, as labor militancy escalates by one unit, trade union membership increases by .966 units. The estimates further suggest that both observed indicators of latent labor militancy are indeed subject to a significant amount of measurement error.

Note that $\gamma_1 > 0$ and $\gamma_2 < 0$ are strong evidence against certain restrictive assumptions employed in the Ashenfelter and Johnson model. Our more general model, however, postulates this very sign pattern under conditions characteristic for the West German economy.

The third section in Table 1 presents five goodness-of-fit statistics, the first three relating to the entire model, the latter two giving information on the two submodels¹⁷. The chi-square statistic provides a probability measure for evaluating the militancy model. It tests the hypothesis that Σ is of the form implied by the model against the alternative that Σ is unconstrained, i.e. that the variables are arbitrarily correlated. We obtain a X^2 of 1.2, which is small for 3 degrees of freedom. Hence, we may conclude that the proposed model appropriately mirrors the causal process that generates the data in the population. The associated p value of 0.75 is the probability of getting a chi-square value larger than 1.2, given that the hypothesized constraints are true. *AGFI* is a goodness of fit index adjusted for degrees of freedom as defined in the notes to Table 1. It is an adjusted measure of the relative amount of the variances and covariances jointly accounted for by the two submodels. It does not depend on sample size, as does the chi-square test, and is reasonably robust against departures from normality. Unfortunately its statistical distribution is not known. However, *AGFI* can and will be used below to compare the fit of different models for the same data. *CDY* is the coefficient of determination for the measurement model. Its value of 0.61 indicates that the observed indicators C and T serve reasonably well as indicators for latent militancy. The coefficient of determination of the structural equations model, *CDS*, is 0.85, which suggests a highly satisfactory fit of this submodel.

In light of the fact that a sample size of 30 is a rather small basis for estimating a LISREL model, the sensitivity of the estimates to variations of sample size was studied. As table 2 demonstrates, cutting off five years from either side of the sample period leaves the results virtually unchanged. This allows us to place some faith in the robustness of our results.

In introducing the main version of the latent militancy model at the beginning of section 2.2., a number of assumptions had to be made which

¹⁷ For a detailed description of these overall goodness-of-fit statistics and several applications thereof see JÖRESKOG and SÖRBOM (1981). Further investigations together with some alternatives are given in BENTLER and BONETT (1980) and BENTLER (1982).

TABLE 2.

CAUSES AND CONSEQUENCES OF LABOR MILITANCY
IN WEST GERMANY; VERSION 1; FULL AND REDUCED SAMPLE PERIODS;
LISREL ESTIMATES. THE NOTES OF TABLE 1 APPLY

	Sample Period		
	1952 - 1981	1952 - 1976	1957 - 1981
β_1	0.742 ** (0.264)	0.740 ** (0.315)	0.732 * (0.288)
γ_1	0.620 ** (0.167)	0.568 ** (0.172)	0.660 ** (0.208)
γ_2	- 0.560 ** (0.163)	- 0.520 ** (0.169)	- 0.540 * (0.199)
γ_3	0.750 ** (0.154)	0.717 ** (0.182)	0.757 ** (0.149)
λ_1	1.000 (c)	1.000 (c)	1.000 (c)
λ_2	0.966 ** (0.306)	0.985 ** (0.332)	0.908 * (0.364)
$\lambda_3, \lambda_4, \lambda_5$	1.000 (c)	1.000 (c)	1.000 (c)
$V[\varepsilon_1]$	0.500 ** (0.172)	0.488 * (0.188)	0.535 * (0.202)
$V[\varepsilon_2]$	0.534 ** (0.176)	0.503 * (0.190)	0.716 ** (0.213)
ψ_1	0.158 ** (0.116)	0.171 (0.132)	0.168 (0.137)
ψ_2	0.443 ** (0.133)	0.499 ** (0.166)	0.334 * (0.121)
χ^2	1.20(3)	1.99(3)	2.39(3)
p	0.75	0.58	0.50
AGFI	0.93	0.85	0.87
CDY	0.61	0.64	0.52
CDS	0.85	0.82	0.86

will now be subjected to closer scrutiny. The more crucial ones concern the choice of indicators for latent militancy and the maximum lag at which the growth rate of real wages enters the formation of rank and file aspirations. Of course, more than two variables could potentially serve as indicators of labour militancy. For one thing, the International Labour Office publishes

two strike series. In addition to the "number of workers involved in industrial conflicts", which has been used in the estimation of our main model, ILO also publishes the "number of labor days lost due to strikes or lock-outs". An attempt to use this series as a third indicator failed, however, as the correlation between the two series proved too strong for certain algorithms used in LISREL. Hines' (1964) work suggests that absolute union membership T may be considered an indicator for labor militancy. Although this indicator appears less plausible than T and has always been highly disputed, the result of introducing T into our model is given as *version 2* in Table 3. λ_{2a} is the factor loading of T and ε_{2a} is the measurement error contained in T . All other symbols have the same meaning as in version 1.

The result is somewhat surprising. On the one hand, T appears to be a significant indicator of M , though with a very high error variance $V(\varepsilon_{2a}) = 0.796$. On the other hand, inclusion of this new indicator does not seem to affect the other coefficients and, hence, the dynamics of the structural equations model to any noticeable degree. This and the fact that goodness-of-fit statistics for the total model are uniformly worse than for version 1 speaks strongly in favor of the latter. This holds although the coefficient of determination for the structural equations model of version 2 is somewhat better than for version 1.

Version 4 starts out from version 1 and includes \dot{r}_{t-2}^e as an additional determinant of rank and file aspirations \dot{r}_t^* . Here almost the opposite effect occurs, as compared to the previous extension of the main model. With a t -value of only 1.36 the coefficient of \dot{r}_{t-2}^e , γ_{1a} , is statistically insignificant, and, again, the other coefficients remain unaffected. Nevertheless, all goodness-of-fit statistics for the total model and for the structural equations model improve. Only the coefficient of determination for the measurement model drops slightly. However, improved fits for the total and the structural equation model have obviously been achieved by including a variable with a statistically insignificant impact. Because of this, and since parsimony is an important virtue considering the size of our sample, we prefer the more compact version 1¹⁸.

The merits of treating labor militancy as a latent variable are clearly brought out by version 3. *Version 3* models the case, which is common in the literature, that militancy is proxied by strike activity only, without

¹⁸ Discrimination between versions 1 and 4 is not crucial anyway, regarding the interpretation and the dynamic behavior of the model. Version 4 gives rise to the second-order equation $(1 - 0.554L + 0.138L^2)r_t = 0.318 U_t^{-2}$. The only noticeable difference to the first order difference equation given in footnote 16 is in a quantitatively unimportant fluctuation introduced by the statistically insignificant coefficient γ_{1a} .

TABLE 3.

CAUSES AND CONSEQUENCES OF LABOR MILITANCY
IN WEST GERMANY; VERSION 2, 3, AND 4; ANNUAL OBSERVATIONS
1952-1981; LISREL ESTIMATES. THE NOTES OF TABLE 1 APPLY

	Version 2	Version 3	Version 4
β_1	0.781 * (0.286)	0.290 (0.156)	0.786 ** (0.277)
γ_1	0.644 ** (0.170)	0.605 ** (0.179)	0.705 ** (0.188)
γ_{1a}			- 1.175 (0.128)
γ_2	- 0.514 ** (0.156)	- 0.610 ** (0.180)	- 0.559 ** (0.160)
γ_3	0.714 ** (0.159)	0.657 ** (0.156)	0.757 ** (0.154)
λ_1	1.000 (c)	1.000 (c)	1.000 (c)
λ_2	0.956 ** (0.331)		0.973 ** (0.309)
λ_{2a}	- 0.687 * (0.318)		
$\lambda_3, \lambda_4, \lambda_5$	1.000 (c)	1.000 (c)	1.000 (c)
$V[\epsilon_1]$	0.568 ** (0.176)		0.519 ** (0.171)
$V[\epsilon_2]$	0.609 ** (0.182)		0.545 ** (0.154)
$V[\epsilon_{2a}]$	0.796 ** (0.219)		
ψ_1	0.093 (0.089)	0.640 ** (0.174)	0.127 (0.103)
ψ_2	0.433 ** (0.129)	0.600 ** (0.136)	0.414 ** (0.128)
χ^2	10.58(7)	5.05(1)	1.74(5)
p	0.16	0.03	0.88
AGFI	0.69	0.71	0.94
CDY	0.51		0.59
CDS	0.90	0.61	0.88

allowing for measurement errors. Though being satisfactory in terms of statistical significance of the estimated coefficients, the fit of this orthodox model is poor and clearly inferior to the *latent* militancy model. This is brought out even more clearly by the squared multiple correlations for the two structural equations, which are 0.69 and 0.57 for the latent militancy model and 0.36 and 0.40 for the model considering strikes only¹⁹.

3. Summary

This investigation moves beyond the status quo of research on labor militancy in two respects. On the theoretical level, a general model of the causes and consequences of labor militancy is proposed. This model includes the strike model by Ashenfelter and Johnson, which has dominated the pertinent literature for more than a decade, as a special case. The need for such a more general model arises out of two sorts of empirical observations noted by other authors and substantiated by our own results²⁰. The first observation is that, more often than not, militancy not only affects real wage growth positively, but also *responds to past real wage growth with a positive coefficient*. The second observation is that *rising unemployment quite often appears to escalate labor militancy* rather than to dampen it. While both these observations are in open contrast to the postulates of the Ashenfelter and Johnson model, they are readily explained within our generalized framework.

On the empirical level, labor militancy is treated as a latent, unobserved variable which affects certain indicators. To the best of our knowledge, no such effort has been published up until now. The estimation results strongly support the proposed model, but, as has already been stressed, strongly contradict the Ashenfelter and Johnson model. It is further demonstrated that the latent militancy model is distinctly superior to the traditional approach, which draws on some strike measure as a single and error-free proxy for labor militancy.

¹⁹ Due to its small sample size properties the usual χ^2 difference test is not applicable (see BENTLER and BONETT (1980) on the small sample size properties of this test).

²⁰ The studies referred to are PALDAM (1981), PALDAM and PEDERSEN (1982) and PHIPPS (1977).

APPENDIX A

The $\Sigma(\theta)$ matrix of the main model (Version 1)

Extensive manipulations applying the rules of covariance algebra transform the equations and assumptions of version 1 of our latent-militancy model into the following variance-covariance structure:

$$\Sigma(\theta) = \begin{array}{ccccc} & \begin{array}{c} \Sigma_{11} \\ \Sigma_{21} \\ \Sigma_{31} \end{array} & & \begin{array}{c} \Sigma_{22} \\ \Sigma_{32} \end{array} & \\ & & & & \begin{array}{c} \Sigma_{33} \end{array} \\ \begin{array}{c} \gamma_1 \phi_{11} + \gamma_2 \phi_{21} \\ \gamma_1 \phi_{11} + \gamma_2 \phi_{22} \end{array} & \begin{array}{c} \lambda_2 \gamma_1 \phi_{11} + \lambda_2 \gamma_2 \phi_{21} \\ \lambda_2 \gamma_2 \phi_{22} + \lambda_2 \gamma_1 \phi_{21} \end{array} & \begin{array}{c} \beta_1 \gamma_1 \phi_{11} + (\beta_1 \gamma_2 + \gamma_3) \phi_{21} \\ \beta_1 \gamma_1 \phi_{21} + (\beta_1 \gamma_2 + \gamma_3) \phi_{22} \end{array} & \begin{array}{c} \phi_{11} \\ \phi_{21} \end{array} & \begin{array}{c} \phi_{22} \end{array} \end{array}$$

with:

$$\begin{aligned} \Sigma_{11} &= \gamma_1^2 \phi_{11} + 2\gamma_1 \gamma_2 \phi_{21} + \gamma_2^2 \phi_{22} + \psi_{11} + \theta_{\epsilon 11} \\ \Sigma_{21} &= \lambda_2 \gamma_2^2 \phi_{11} + 2\lambda_2 \gamma_1 \gamma_2 \phi_{21} + \lambda_2 \gamma_2 \phi_{22} + \lambda_2 \psi_{11} + \theta_{\epsilon 31} \\ \Sigma_{31} &= \beta_1 \gamma_1^2 \phi_{11} + (1 + 2\beta_1) \gamma_1 \gamma_3 \phi_{21} + \gamma_2 (\beta_1 \gamma_2 + \gamma_3) \phi_{22} + \beta_1 \psi_{11} + \theta_{\epsilon 31} \\ \Sigma_{22} &= \lambda_2^2 \gamma_1^2 \phi_{11} + 2\lambda_2^2 \gamma_1 \gamma_2 \phi_{21} + \gamma_2 \phi_{22} + \psi_{11} + \theta_{\epsilon 22} \\ \Sigma_{32} &= \lambda_2 \beta_1 \gamma_1^2 \phi_{11} + (1 + 2\beta_1) \lambda_2 \gamma_1 \gamma_3 \phi_{21} + \lambda_2 \gamma_3 (\beta_1 \gamma_2 + \gamma_3) \phi_{22} + \lambda_2 \beta_1 \psi_{11} + \theta_{\epsilon 32} \\ \Sigma_{33} &= \beta_1^2 \gamma_1 \phi_{11} + 2\gamma_1 \beta_1 (\beta_1 \gamma_2 + \gamma_3) \phi_{21} + [\beta_1 (\beta_1 \gamma_2^2 + 2\gamma_2 \gamma_3) + \gamma_3^2] \phi_{22} + \\ &+ \beta_1^2 \psi_{11} + \psi_{22} \end{aligned}$$

All elements of Σ draw on elements of Λ_y , Λ_x , B , Γ , Φ , Ψ , θ_ϵ and/or θ_δ only. LISREL (Jöreskog and Sörbom, 1981) minimizes the fitting-function $F(\theta)$, as given in section 2.1., by an iterative procedure which starts at initial estimates $\theta^{(1)}$, arrived at by unweighted least squares, and successively generates new points $\theta^{(2)}$, $\theta^{(3)}$, ... in the parameter space such that $F[\theta^{(t+1)}] < F[\theta^{(t)}]$ until convergence is obtained. The minimization procedure uses first-order derivatives and approximations to the second-order derivatives of F (information matrix)^{21 22}.

²¹ The information matrix has to be positive definite in order to calculate the standard errors of the estimations. It can be shown that this is almost always the case for fully identified models (JÖRESKOG and SÖRBOM, 1981). Hence, positive definiteness of the information matrix is a useful criterion for checking the identification status of a model, whenever it is difficult or too involved to resolve it formally.

²² See, e.g. JÖRESKOG (1978) or STAPLETON (1979).

APPENDIX B

The Data

1. Sources:

1.1. *Statistisches Jahrbuch der Bundesrepublik Deutschland*, Wiesbaden, various years.

\dot{w}^n = Percentage rate of change of negotiated hourly wage rates for males minus consumer-price inflation rate.

\dot{w}^e = Percentage rate of change of effective hourly earnings in industry minus consumer-price inflation rate.

U = Unemployment rate.

T = Number of workers organized in trade unions.

1.2. *International Labor Office Yearbook*, various years.

S = Natural logarithm of the number of workers involved in strikes or lockouts (alternative estimates using man-days lost due to strikes or lockouts are also discussed in the text).

2. Notes:

2.1. Percentage rates of change are proxied by first differences of natural logs.

2.2. Source 1.1. gives end-of-year numbers for T . Midyear numbers are constructed by linear interpolation.

2.3. $\dot{T}_t = T_t - T_{t-1}$

2.4. Strike data are transformed to natural logs since original data exhibit a strongly skewed distribution which pronouncedly differs from normality.

REFERENCES

- ASHENFELTER O. and JOHNSON G.E., "Bargaining Theory, Trade Unions, and Industrial Strike Activity", *American Economic Review*, 1969, 59, 35-49.
- BENTLER P.M., "Linear Systems with Multiple Levels and Types of Latent Variables", in K.G. Jöreskog and H. Wold, eds., *Systems under Indirect Observation: Causality - Structure - Prediction*, Amsterdam & New York: North-Holland, 1982.
- and BONETT D.G., "Significance Tests on Goodness-of-Fit in the Analysis of Covariance Structures", *Psychological Bulletin*, 1980, 88, 588-606.
- DICKS-MIREAUX L.A. and DOW J.C.R., "The Determinants of Wage Inflation: United Kingdom, 1946-1956", *Journal of the Royal Statistical Society*, 1959, Series A 122, 145 ff.

- EDWARDS P.K., "Methodological Issues in Quantitative Strike Analysis", *Industrial Relations*, 1979, 18, 17-24.
- GÄRTNER M., "Politik und Arbeitsmarkt: Eine Uebersicht über ausgewählte Makrotheorien", *Zeitschrift für die gesamte Staatswissenschaft*, 1981, 137, 252-83.
- , "Causality Patterns in Aggregate Labor Markets: A Statistical Analysis of West German Data", *Empirical Economics*, 1983, 8, 229-45.
- , "Strikes and the Real Wage-Employment Nexus: A Hicksian Analysis of Industrial Disputes and Pay", *Journal of Labor Research*, 1985, 6, 323-36.
- HICKS J., *The Theory of Wages*, London: MacMillan, 1932 (2nd Edition 1963).
- HINES A.G., "Trade Unions and Wage Inflation in the United Kingdom 1893-1961", *Review of Economic Studies*, 1964, 31, 221-52.
- JÖRESKOG K.G., "A General Approach to Confirmatory Maximum Likelihood Factor Analysis", in K.G. Jöreskog and D. Sörbom, eds., *Advances in Factor Analysis and Structural Equation Models*, Cambridge, Mass.: Abt Books, 1979.
- , "Analysis of Covariance Structures", *Scandinavian Journal of Statistics*, 1981, 8, 65-92.
- and SÖRBOM D., *LISREL-Analysis of Linear Structural Relations by the Method of Maximum Likelihood*, University of Uppsala, 1981.
- PALDAM M., "Industrial Conflict and the Phillips-Curve. An International Perspective", Memo 1980-4 (revised June 1981), Institute of Economics, University of Aarhus.
- and PEDERSEN P.J., "The Macroeconomic Strike Model: A Study of Seventeen Countries, 1948-1975", *Industrial and Labor Relations Review*, 1982, 35, 504-21.
- PHIPPS A.J., "Strike Activity in Australia", *Economic Record*, 1977, 53, 297-319.
- PURDY D.L. and ZIS G., "On the Concept and Measurement of Trade Union Militancy", in D. Laidler and D.L. Purdy, eds., *Inflation and Labour Markets*, Manchester: University Press, 1974, 61-78.
- ROSS A.M., *Trade Union Wage Policy*, Berkeley and Los Angeles: University Press, 1948.
- STAPLETON D.C., "Analyzing Political Participation Data with a MIMIC-Model", in K. Schuessler, ed., *Sociological Methodology*, San Francisco, 1978.
- STERN R.N., "Methodological Issues in Strike Analysis", *Industrial Relations*, 1978, 17, 32-42.
- ZEUTHEN F., *Problems of Monopoly and Economic Welfare*, London: MacMillan, 1930.

CAUSE E CONSEGUENZE DELLA MILITANZA SINDACALE. IL FATTORE COMUNE NEL CONFLITTO INDUSTRIALE E NELLA DINAMICA DELLA SINDACALIZZAZIONE

Questa indagine si muove oltre lo *status quo* della ricerca sulla militanza

sindacale sotto due aspetti. Al livello teoretico viene qui proposto un modello generale delle cause e delle conseguenze della militanza sindacale. Il modello qui discusso ingloba come caso speciale il modello dello sciopero dovuto a Ashenfelter e Johnson, modello che ha dominato la letteratura rilevante per più di un decennio. Il bisogno di passare a un nuovo modello nasce da un duplice tipo di osservazione empirica, sostanziata anche attraverso i risultati di questo lavoro. La prima osservazione è che, molto spesso, non soltanto la militanza influisce positivamente sulla crescita salariale in termini reali, ma risponde con coefficiente positivo alla crescita passata del salario reale. La seconda osservazione è che la disoccupazione crescente è spesso un fattore di stimolo della militanza sindacale piuttosto che un fattore del suo contenimento. Entrambe queste osservazioni si trovano a essere in aperto contrasto coi postulati di Ashenfelter e Johnson, mentre appaiono di agevole spiegazione nella struttura generale proposta dal modello qui discusso.

Empiricamente, la militanza sindacale viene trattata quale variabile latente non osservata che influenza certi indicatori. A conoscenza degli autori si tratta di un'impostazione nuova: essa trova conforto in soddisfacenti risultati empirici, che altri modelli non sono in grado di spiegare. Il modello della militanza sindacale appare anche superiore all'impostazione tradizionale, la quale usa certe misure relative allo sciopero quale *proxy* della militanza.

DIRECT AND INDIRECT LABOR AND DEMAND FOR WORKERS AND WORKING HOURS

by
PAOLO PALAZZI * and PAOLO PIACENTINI *

Abstract

The main purpose of the paper is to estimate a function of demand for labor (both for production workers and working hours) for the U.S. Manufacturing Industry for the period 1956-84.

The proposed functional form includes, among explanatory variables, the stock of capital and its average age in addition to output and trend. Such a specification is aimed at capturing the incidence of the "indirect" component of labor input, whose behavior cannot be described in a flow-to-flow relationship with output, and is assumed to be complementary to the accumulation of capital.

The average age variable is aimed at partially capturing the effect of the embodied technical progress.

The results of the estimations appear to be encouraging for the basic hypotheses.

Introduction

In recent years "segmentation" theorists have increasingly challenged the positing of homogeneous labor as a feasible model. Different characteristics of two (or more) segments of the labor force are often expressed in terms of social, behavioral, racial factors, etc.; these elements are essentially extrinsic to the strictly technical features of the production process. This paper

* University of Rome.

explores the implications of a particular type of heterogeneity in the work force which stems from differences in the role and characteristics of labor utilization *inside* the production process. The most obvious example is the distinction between white collar and manual labor; but also within the aggregate which is normally classified in statistical sources as "production" or "operative" labor, it seems increasingly relevant to make a distinction between "directly" and "indirectly" productive components. This dichotomy is not expressly related to the personal characteristics of the labor force (such as sex or age referred to in segmentation theories); rather, it emerges from differences in the nature of the tasks performed, independently of these individual factors. In this paper various implications of this duality hypothesis will be developed in terms of a model of labor demand; major results of empirical estimates for manufacturing in the United States from 1956 to 1984 will also be outlined.

The Composition of Total Labor Input

The traditional approach to empirical analysis of demand for operative labor posits a homogeneous labor input whose relationship to the level of production (in terms of elasticity, lagged response, etc.) is explored using different behavioral hypotheses to explain demand for labor. Introducing a concept of "labor hoarding" and/or existence of lags in the adjustment of current input requirements to optimal demand does not contradict the core of the traditional vision of labor input as flow of productive services functionally related to the flow of production.

In partial contrast to this framework, it will be proposed that it is possible and useful to distinguish between "direct" and "indirect" components within the aggregate of "production" labor. This distinction is to a certain degree an extension of the traditional dichotomy between productive and "overhead" labor, the latter generally defined as clerical, administrative and a part of technical personnel. It is our belief that a significant sector exists, within the aggregate designated as "production" or "operative" labor, whose functional determination and dynamic behavior cannot be incorporated into the "flow against flow" theory; this sector's "quasi-fixed" nature in the face of short-run output fluctuation calls for further or alternative causal explanation. Unfortunately, the distinction between "directly productive" and "indirect" labor cannot be quantitatively ascertained with available statistics; rigorous investigation would require case studies of individual processes. Furthermore, over the course of a work year or even during a work day, the

same personnel may perform both direct and indirect tasks; thus, a clear identification of direct versus indirect workers and corresponding work-time may not be feasible. The introduction of this dichotomy into labor input theory, may, nevertheless, be useful in interpreting current behavior of aggregate work hours and employment.

In particular, this new aspect may bear upon the perennial controversy over the interpretation of empirical evidence of increasing returns to labor in the short-run, and possible ways of reconciling this evidence with orthodox models of factor demand by firms.

According to the theory of the influence of an "overhead" component upon behavior of aggregate labor input, this component's rigidity throughout the cycle would explain the procyclical behavior of measured (average) productivity; such a hypothesis has often been advanced as the most convincing explanation of the "increasing returns" paradox. As a rule, however, the "overhead" component has been wholly identified with nonproduction (clerical/administrative/certain technical) workers; while the empirical analysis restricted to the aggregate of production workers, still prevalently showed evidence of SRIRL¹. It is believed that this framework, to wit searching for a rationale for non-proportionalities between work hours and output variability within the production process itself, might represent a valid groundwork for interpreting these phenomena.

The Model

Let those work hours, functionally related to output on a flow-to-flow basis, be termed "directly productive (or more briefly, direct) hours". Any variation of production thus requires some variation of direct hour input.

Indirectly productive (indirect) hours are those work hours that, *although technologically necessary for the realization of the production process, do not depend on output flow.*

All production processes, even the most primitive, require a quota of total labor time for overseeing the instrumental inputs and the goods involved in the production process (maintenance and repair, storage, etc.).

Apart from specifying indirect tasks, however, it is more important to point out that advances in technological sophistication have a tendency to raise the degree of separation of the worker from the production process in its narrower sense. The cutoff point is automation, where the entire process

¹ For a recent discussion, see COSTRELL (1982). For evidence of increasing returns when production workers alone are considered, see HULTGREN (1965), FAIR (1969).

of transformation of material inputs is accomplished by automata; residual hours of operative workers may be considered a complement to machinery, necessary for its proper operation and control.

A simple formalization of an empirically testable model will now be attempted on the basis of these definitions.

a) "Direct" hours imply a flow-to-flow relation with output; an elasticity coefficient specifies the required proportionality. At this point what interests us is a theoretical formulation of an optimal hours requirement; therefore, specifications for lagged response or "labor excess" with respect to an optimal level, due to the existence of "labor hoarding" margins, will not be introduced.

The relationship between hours and output will be altered over time by technical progress. Without imposing unnecessary restrictions, one may imagine technical progress operating through both a "disembodied" trend and an "embodied" component; this advancement necessitates prior investment to incorporate new technology and to improve average efficiency of capital stock.

Thus the functional relationship for direct hours may be specified in the following form:

$$HP_t = \frac{e^{-\delta t} Y_t^\alpha}{\mu(K_t)}$$

where HP = direct hours, δ is the exogenous rate of disembodied technical progress and $\mu(K_t)$ is an efficiency index, to be specified, associated with capital stock in operation during period t .

b) Detailed specification of the indirect hours requirement would call for an investigation of the technical structure of each particular process. No such information is available at the level of aggregate data; thus, a plausible approximation must be determined. If indirect hours are chiefly devoted to servicing productive equipment rather than to manufacturing of materials, then the capital stock in use could be considered an indicator of productive capacity; the indirect hours requirement is derived from this reasoning. In such a simple hypothesis, the level of these hours is related to the capital stock in operation and is "fixed" in the short-run with respect to variations in output. It should be noted that the relationship implied between labor and capital is reversed with respect to a traditional two-factor production function; this is so because an assumption of complementarity is introduced between indirect components of labor input and capital stock.

The following is a quote from A.M. Okun:

"For example, labor may be needed to maintain overhead capital: a worker may be needed merely to lubricate an idle machine. In that case, the employment of such a worker is necessary in a slump, although the reason he is needed stems not from any fixity in the labor factor, but rather from a production function in which capital and labor are complementary (and not substitutive) in the short-run" (Okun, 1981, p. 17).

In the hypothesis set forth in this paper short-run complementarity concerns a section of the total work force; it is agreed that the requirement for indirect labor to service capital stock should be mainly related to the stock of the latter, rather than to its rate of utilization².

Labor saving through technical progress should also occur with respect to indirect hours through disembodied and embodied (e.g., new machinery requiring less servicing) components. Hence, the equation would be as follows:

$$HF_t = \frac{e^{-\delta' t} K_t^\beta}{\mu'(K_t)}$$

where HF = indirect hours, and K = capital stock.

An appropriate operational model for empirical estimates must be specified in order to test the hypothesis of a dual structure of labor input.

The dependent variable in the equation shall be represented by the sum of HP and HF , to wit total labor input (number of hours worked by "production" workers, or the number of production workers employed); the dependent variable is specified in this way given the impossibility, as noted earlier, of making separate estimates of "directly" and "indirectly" productive labor input.

Formulation of the function $\mu(K_t)$, indicating the degree of efficiency embodied in capital stock, is surely the most difficult task. An "efficiency index" should include the rate of "embodied" technical progress and some indicator of the vintage structure of the stock in use. From the obvious consideration that recent investment in all likelihood incorporate more efficient technologies, a higher share of newer to total capital stock should increase efficiency, *coeteris paribus*. For the rate of embodied technical progress, we are here only able to add a further trend element, while we assume that the temporal structure of the stock in use can be adequately described by the average age of the stock; i.e. greater average age is a proxy for lower efficiency. Thus the formulation is as follows:

² Complementarity between total labor input and capital stock also emerges from estimations in pluri-factor (KLEM) input demand models. (See MORRISON and BERNDT, 1981).

$$\mu(K_t) = \mu'(K_t) = \lambda_0 e^{\lambda t} A_t^{-\gamma}$$

where λ = rate of "embodied" technical progress; A_t = average age of capital stock at time t .

The complete function for total labor input requirement thus becomes:

$$(1) \quad H_t = HP_t + HF_t = (e^{-\delta t} Y_t^\alpha + e^{-\delta t} K_t^\beta) \lambda_0 e^{-\lambda t} A_t^\gamma$$

The following log linear expression for the empirical estimation may be derived from this function:

$$(1') \quad \lg H_t = C + \alpha(1-f) \cdot \lg Y_t + \beta f \lg K_t + \\ + \gamma \lg A_t - [\lambda + \delta'f + \delta(1-f)]t$$

where C = constant term; $f = \frac{HF}{H}$ (the share of indirect labor to total labor)³.

It is clear from the above equation that the coefficients of output and capital stock are determined by the elasticities of direct and indirect labor weighted by their respective share of total labor input. Hence, the empirical estimate of labor elasticity with respect to output will be decisively influenced by labor input structure.

Constant or decreasing returns on directly productive labor input become compatible with apparent evidence of increasing returns when total labor input is considered.

Empirical Results

Equation (1') and a modified version were tested on U.S. manufactur-

³ The linear function expressed in logarithms may be considered to be approximately hold, given the expression in terms of variation rates:

$$\dot{H} = \frac{HP}{H} \dot{HP} + \frac{HF}{H} \dot{HF} = -[\lambda + \delta'f + \delta(1-f)] + \alpha(1-f) \dot{Y} + \\ + \beta f \dot{K} + \gamma \dot{A}$$

Unfortunately, some "external" information would be required for the purpose of knowing the share " f " and thus identifying elasticity parameters. Patterns of technical progress leading to changing proportions of direct and indirect labor inputs may in the longer run imply instability of the estimated elasticities.

ing data for the period 1956-1984, for both total work hours and number of production workers ⁴.

The modified version of equation (1) is derived from a simple algebraic operation which allowed introduction of the capital/output ratio as a variable (K_t/Y_t). Equation (1) can be rewritten in the following way:

$$(2) \quad H_t = \left[e^{-\delta t} Y_t^\alpha + e^{-\delta' t} \left(\frac{K_t}{Y_t} \right)^\beta Y_t^\beta \right] \cdot A_t^\gamma e^{\lambda t}$$

And the following logarithmic expression may be derived from the above equation:

$$(2') \quad \lg H_t = C + [\alpha(1-f) + \beta f] \lg Y_t + \\ + f\beta \lg \left(\frac{K_t}{Y_t} \right) - [\lambda + \delta f + \delta'(1-f)] t + \gamma \lg A_t$$

The only difference between equations (1') and (2') lies in the value of the production coefficient; in equation (2') the value is determined by the sum of the production and capital stock coefficients of equation (1').

The two equations have been formulated utilizing both the number of hours worked by production workers and, directly, the number of production workers in the manufacturing sector, to determine labor input. Manufacturing production data was used to determine the value of output. With regard to capital stock for absolute accuracy only net capital data should be used, in view of the fact that this data more accurately represents the make-up and technological structure of machinery actually in use; as is well-known, however, the depreciation computation in the available data is based on very elementary (straight line) depreciation hypotheses; for this reason, both series of net capital and gross capital stock were used.

Estimates were formulated utilizing annual data; in this way structural components of labor demand, independent of the short-run cycle, may be more accurately determined. The estimates are expressed in both a logarithmic base and in terms of variation rates; the latter are suitable for verifying the stability of the relationships and for reducing possible phenomena of multicollinearity. The results of our estimates are indicated in Table 1 and

⁴ Sources: Average weekly hours and production indexes — *Economic Report of the President*, February 1985 - Tables B-38 and B-42.

Production and related workers. U.S. Dept. of Labor. *Employment and Earnings*, 1984.

Constant cost valuation of fixed nonresidential private capital (gross and net) and average age (gross and net) — U.S. Dept. of Commerce. Bureau of Economic Analysis. *Fixed Reproducible Tangible Wealth in the U.S. 1925-84* - Table A-1.

2. It seems feasible to state in general that the results tend to confirm the initial hypotheses with respect to labor input expressed in work hours as well as to the number of production workers ⁵.

TABLE 1.
DEPENDENT VARIABLE: TOTAL HOURS WORKED BY PRODUCTION WORKERS

Trasformations	C	<i>t</i>	Y	KN	KG	KN/Y	KG/Y	A_{t-1}	R ² C	DW
<i>lg</i>	- 1.63 (2.27)	- 0.05 (16.21)	0.84 (39.9)	0.70 (6.76)				0.57 (5.00)	0.986	1.182
<i>lg</i> - CO	- 1.59 (1.68)	- 0.05 (12.10)	0.86 (32.9)	0.68 (4.75)				0.55 (3.76)	0.999	1.831
<i>lg</i>	- 4.85 (4.07)	- 0.05 (16.21)	1.54 (14.27)			0.70 (6.76)		0.57 (5.00)	0.986	1.182
<i>lg</i> - CO	- 4.73 (2.95)	- 0.05 (12.10)	1.54 (10.78)			0.68 (4.65)		0.55 (3.76)	0.999	1.931
<i>lg</i>	- 4.03 (4.16)	- 0.06 (16.37)	0.91 (45.26)		0.94 (7.89)			0.63 (4.7)	0.990	1.258
<i>lg</i>	- 8.34 (5.52)	- 0.06 (16.37)	1.85 (14.37)				0.94 (7.89)	0.63 (4.7)	0.990	1.258
RV	- 4.49 (11.16)		0.87 (25.37)	0.52 (3.36)				0.37 (2.09)	0.976	2.408
RV	- 4.91 (10.33)		1.46 (9.82)			0.60 (3.65)		0.40 (2.31)	0.977	2.182
RV	- 5.6 (9.15)		0.88 (28.79)		0.90 (3.85)			0.63 (2.29)	0.978	2.535
RV	- 6.05 (8.35)		1.79 (7.86)				0.90 (3.82)	0.55 (2.11)	0.978	2.079

Legenda: C = Constant; *t* = Trend; Y = Production index; KN = Capital stock net; KG = Capital stock gross; A = Average age of capital stock (net or gross); *lg* = Natural logarithm; RV = Yearly rate of variation; CO = Cochrane-Orcutt transformation.

⁵ For logarithmic estimates where the D.W. index indicated the presence of first order autocorrelation, estimates on variables, transformed in accordance with the Cochrane-Orcutt method, are also included. (Note how the coefficient values show good stability).

TABLE 2.

DEPENDENT VARIABLE: NUMBER OF PRODUCTION WORKERS

Transformations	C	t	Y	KN	KG	KN/Y	KG/Y	A_{t-1}	R^2C	DW
lg	- 4.8 (6.12)	- 0.05 (14.36)	0.71 (31.09)	0.76 (6.76)				0.41 (3.31)	0.980	1.377
$lg - CO$	- 5.33 (5.4)	- 0.05 (11.63)	0.71 (27.8)	0.85 (5.74)				0.47 (3.17)	0.992	1.820
lg	- 8.31 (6.4)	- 0.05 (14.36)	1.47 (12.55)			0.76 (6.76)		0.41 (3.31)	0.980	1.377
$lg - CO$	- 9.22 (5.56)	- 0.05 (11.63)	1.56 (10.41)			0.85 (5.74)		0.47 (3.17)	0.992	1.820
lg	- 4.8 (3.3)	- 0.05 (9.58)	0.79 (26.4)		0.73 (4.11)			0.07 (.36)	0.972	1.237
lg	- 8.16 (3.61)	- 0.05 (9.58)	1.52 (7.93)				0.73 (4.11)	0.07 (.36)	0.972	1.237
RV	- 4.85 (9.86)		0.66 (15.85)	0.89 (4.67)				0.53 (2.45)	0.948	2.307
RV	- 5.22 (8.09)		1.55 (7.68)			0.88 (3.97)		0.46 (1.97)	0.940	1.988
RV	- 6.38 (7.76)		0.70 (16.99)		1.36 (4.33)			0.74 (2.02)	0.944	2.125
RV	- 6.19 (5.6)		1.78 (5.14)				1.07 (2.96)	0.38 (.96)	0.927	1.92

Legenda: C = Constant; t = Trend; Y = Production index; KN = Capital stock net; KG = Capital stock gross; A = Average age of capital stock (net or gross); lg = Natural logarithm; RV = Yearly rate of variation; CO = Cochrane-Orcutt transformation.

In particular, the confirmation of the existence of a direct relationship between labor demand for production workers and hours and industrial productive capacity, indicated by (net or gross) capital stock and by the capital/income ratio, seems significant. This demonstrates how the theory of indirect labor also existing within the aggregate of production workers can be tested using variables representing technological structure. Accepting this

framework, one can draw interesting inferences regarding labor elasticity coefficients with respect to production. In all our estimates, the coefficients were found to be less than unity, demonstrating increasing returns. As shown earlier, however, this coefficient is the product of directly productive labor elasticity and its weight in aggregate labor input. For example, these results can be considered compatible with a coefficient of actual elasticity of direct work hours with respect to output equal to or greater than one, where the share of indirect work hours is greater than a value of 12-13%.

As was to be expected, the production coefficient in the employment equations is considerably lower. With regard to labor demand expressed in number of workers, institutional rigidity factors (adjustment costs, etc.) should be considered greater here than in the case of work hours; hence, elasticity coefficients equal to or greater than unity, also with regard to directly productive workers only, are extremely unlikely. Finally, it should be noted how the logarithmic estimates of labor demand, in which the capital/income ratio variable is clearly stated, demonstrate that the coefficients correspond almost exactly to the algebraic sum of the production and capital coefficients in equation (1')⁶; in coherence with the algebraic transformation implied.

All coefficients of labor elasticity demonstrate a significant direct relationship to capital stock variation. Coefficient values are considerably lower for estimates in which net capital is used; a reasonable explanation, however, may be found for this systematic divergence⁷.

It should be noted that elasticity coefficients, with respect to capital, are higher for labor input in terms of workers than in terms of work hours; this is in contrast with what occurred for production. As could be predicted, however, the value vary only slightly, as they refer to fixed labor input in which work hours and employment should have the same dynamic.

Average age coefficients are positive and above significance levels, with a time lag of one year in all cases, with the exception of the equations

⁶ In the variation rate estimates, the production coefficient in equation with the capital/product ratio also has values close to the sum of the Y and K coefficients in the base equation; in this case, however, the algebraic construction does not entail an exact relationship.

⁷ Defining labor input elasticity in relation to net and gross capital as ε_N and ε_G respectively, one can easily demonstrate the following equation, approximating in terms of finite variation:

$$\varepsilon_G/\varepsilon_N = \frac{1 - \Delta D/\Delta K_G}{1 - D/K_G}$$

where D = cumulative depreciation ($K_G = K_N + D$)

If investments are prevalently increasing over time, the marginal incidence of D is less than the average incidence; thus $\varepsilon_G/\varepsilon_N$ will be greater than one. This is also confirmed by the regression coefficients for variation rates of K_G over K_N .

using gross capital stock in employment equations. The existence of lag in the significance of the age variable should reflect, in our opinion, a lag in full activation of current investments, for which the structure of the capital stock one-period before is more effective as indicator of the capacity in use. The hypothesis that the efficiency component can be incorporated through a synthetic indicator of the time structure of investments is thus confirmed.

The value of the trend variable summarizes the labor saving effects of both embodied and disembodied technical progress; these are highly negative, indicating that demand for labor input is reduced, *coeteris paribus*, by approximately 5% annually in terms of hours and by slightly less in terms of number of workers⁸.

Further study may lead to a better specification of the characteristics of the technology in the model. The "average age of capital" variable, apart from its dependence on conventional hypothesis of the lifespan of equipment and plant, obviously covers only a "moment" in the time distribution of productive machinery. The high values and significance of the trend reflect formulation limits of the model, i.e. the inability of introducing a specification of further variables and factors influencing the pace and the shape of technical progress⁹.

It seems possible to affirm, nonetheless, that, overall, the results are encouraging for our basic hypothesis. In our opinion, the functional distinction in labor input between directly and indirectly productive components with respect to production workers can constitute a useful approach to the debate around the problem of the dynamic of labor returns. In particular, it is our belief that the introduction of this distinction – for labor input expressed both in terms of number of workers and in terms of work hours – can represent a plausible answer to the difficulties encountered in interpreting empirical estimates of labor returns.

REFERENCES

- COSTRELL R.M., "Overhead Labor and the Cyclical Behavior of Productivity and Real Wages", *Journal of Post-Keynesian Economics*, Winter 1981/82, 277-90.
- FAIR R.C., *The Short-Run Demand for Workers and Hours*, Amsterdam: North-Holland, 1969.

⁸ In the equations in terms of annual variation rates, the share of labor saved can be measured by the negative value of the constant.

⁹ Further study may lead to a better specification of the characteristics of the technology of the model, for example, by testing different depreciation hypotheses.

- HULTGREN T., *Cost, Prices and Profits: Their Cyclical Relations*, New York: National Bureau of Economic Research, 1965.
- KUH E., "Cyclical and Secular Labor Productivity in United States Manufacturing", *Review of Economics and Statistics*, 1-2/1965, 1-12.
- MORRISON C.J., BERNDT E.R., "Short-Run Labor Productivity in a Dynamic Model", *Journal of Econometrics*, 3/1981, 339-65.
- OI W.Y., "Labor as a Quasi-Fixed Factor", *Journal of Political Economy*, Dec. 1962, 538-55.
- OKUN A.M., *Prices and Quantities: A Macroeconomic Analysis*, Oxford: Basil Blackwell, 1981.
- PALAZZI P., PIACENTINI P., "Investimenti, stock di capitale e domanda di lavoro nell'industria italiana", *Economia e Lavoro*, 4/1982, 67-76.
- SIMS C.A., "Output and Labor Input in Manufacturing", *Brookings Papers on Economic Activity*, 3/1974, 695-735.
- SOLIGO R., "The Short-Run Relationship between Employment and Output", *Yale Economic Essays*, Spring 1966, 160-215.

LAVORO DIRETTO E INDIRETTO IN UN MODELLO DI DOMANDA DI LAVORO

Il saggio stima, per l'industria manifatturiera americana e per il periodo 1956-84, una funzione di domanda di lavoro, espressa sia in termini di ore totali lavorate che di occupazione operaia.

L'aspetto che differenzia la forma funzionale adottata dalle equazioni più convenzionali è rappresentato dall'introduzione, fra le variabili esplicative, dello stock di capitale e della sua età media in addizione al volume della produzione. Nel modello proposto, la relazione fra input di lavoro e accumulazione di capitale ha segno positivo, in quanto si ritiene che componenti di lavoro "indiretto" si muovano in modo complementare allo stock di attrezzature produttive a cui sono asserviti. Il lavoro indiretto viene definito come quello che, sebbene tecnologicamente indispensabile alla realizzazione del processo produttivo, non dipende nel breve periodo dal flusso di produzione attivata. L'introduzione dell'età media si propone lo scopo di catturare almeno in parte l'incidenza del progresso tecnico "incorporato" sulla produttività e quindi sui fabbisogni di lavoro. Un ciclo favorevole di investimenti, che abbassa l'età media dello stock in uso, dovrebbe quindi, *coeteris paribus*, abbassare la domanda di lavoro.

I risultati della stima sono nel complesso confortanti per le ipotesi di base del modello, e riteniamo possano rappresentare un punto di riferimento per il dibattito circa l'interpretazione della dinamica dei rendimenti del lavoro nel breve periodo.

LA DIMENSIONE AZIENDALE NELLA TEORIA DELL'IMPRESA

di

ROBERTO BUSON *

1. *Considerazioni introduttive*

Gli economisti insistono sempre nel trattare l'impresa come una sorta di dato teorico, di postulato perennemente uguale, in sé non modificabile, soggetto, più che altro per destino naturale, a una unica condotta razionale.

Anche quando più necessitava, per evidenti esigenze analitiche, di svincolarsi da certi presupposti astratti, l'antinomia tra l'impresa reale e quella teorica non solo permaneva, ma si accentuava.

Ora non è chi non veda che per certi problemi economici l'esistenza dell'impresa, come dato del reale, non può essere affatto ignorata. Molto opportunamente Machlup (1980, p. 40) ha messo in rilievo che, se l'oggetto dell'analisi è la distribuzione delle imprese secondo la loro dimensione o se invece è lo sviluppo o ancora l'organizzazione, allora diventa indispensabile insistere « su un alto grado di corrispondenza tra l'oggetto pensato e l'oggetto osservato ».

Questo lavoro, che mira allo studio della variabile dimensionale, si inserisce in quest'ottica. La dimensione infatti acquista un diverso significato a seconda che venga letta come processo o come stato. Nel primo caso l'impresa, e di conseguenza la dimensione, viene colta nel suo divenire: come la risultante cioè di un processo dinamico di integrazione e di adattamento alla struttura che la circonda. Quindi come capacità, da parte dell'impresa stessa, di governare le variabili endogene di fronte a un mercato e a un ambiente in continua evoluzione. Nel secondo caso invece l'impresa viene a specificare le potenzialità connaturate a una determinata scala produttiva.

* Università di Padova, Facoltà di Giurisprudenza.

Viene cioè a individuare meri rapporti quantitativi, che sono accessibili a quell'impresa e non ad altre.

La tesi principale di questo studio è che la dimensione in sé poco o nulla può dire sulla maggiore o minore efficienza di una unità produttiva. E che il teorema di Marshall secondo cui molti vantaggi della divisione del lavoro possono essere conseguiti da piccole fabbriche purché ve ne sia un numero elevato nella stessa attività, costituisce il nucleo di una più generale teoria in tema di rapporti fra imprese.

Nel corso del saggio, per meglio sviluppare queste considerazioni: (1) cercheremo di individuare la struttura dimensionale dell'industria manifatturiera italiana alla luce dell'ultimo censimento industriale; (2) verificheremo se la rielaborazione di Marshall del noto principio smithiano, secondo cui la divisione del lavoro è limitata dall'ampiezza del mercato, sia sostanzialmente un'anticipazione delle nuove teorie in tema di decentramento produttivo.

2. *L'evidenza empirica*

Per prima cosa delineeremo il processo evolutivo che ha caratterizzato la struttura dimensionale dell'industria manifatturiera italiana nell'arco di tempo che va tra i due ultimi censimenti industriali (1971-1981). Ai fini che ci proponiamo prenderemo dapprima in considerazione l'andamento, sia a livello aggregato che disaggregato, dei valori relativi alla distribuzione delle unità locali e dell'occupazione per classi di ampiezza dei rispettivi addetti; quindi la dinamica della dimensione media per settore produttivo.

I dati su cui soffermeremo inizialmente le nostre analisi sono riportati nella Tab. 1. Dalle indicazioni che emergono si può facilmente osservare che nel decennio 1971-1981 la struttura dell'industria manifatturiera italiana ha subito notevoli mutamenti. Innanzitutto si è assistito a un consistente allargamento della base produttiva: l'incremento degli stabilimenti è stato pari al 21%, corrispondente a un saldo netto di 107.727 unità. Tuttavia il processo di crescita non ha investito in egual misura le varie classi dimensionali. Accanto a una crescita modesta delle unità con uno o due addetti, ha fatto riscontro una considerevole espansione delle unità locali con un numero di occupati compresi tra 3 e 50¹. In questa fascia dimensionale gli stabilimenti

¹ Anche negli Stati Uniti il peso delle piccole dimensioni aziendali è rilevante. Come ricorda MARTELLARO (1982, p. 553), in un suo recente articolo, le piccole imprese contano circa il 43% del Prodotto Nazionale Lordo e il 50% dell'occupazione. Più precisamente « Of some 12 million private business in the United States today, 10.5 million (87.5 per cent) are not-corporate firms. Five million are classified as "small business" – 95 per cent of them have fewer than 20 employees. Only 0.7 per cent have more than 100 employees »

TAB. 1 - Distribuzione degli stabilimenti e dell'occupazione secondo la dimensione misurata dal numero di addetti. Periodo 1971-1981.

Classi Dimensionali	UNITÀ LOCALI					OCCUPAZIONI				
	1971	Quota %	1981	Quota %	Variazioni 1981-1971 %	1971	Quota %	1981	Quota %	Variazioni 1981-1971 %
Fino a 2	310.013	60,36	352.780	56,79	13,8	401.673	7,85	455.737	7,95	13,5
3 - 5	93.027	18,61	115.283	18,55	23,9	345.456	6,75	432.823	7,55	25,3
6 - 9	40.135	7,81	58.099	9,35	44,8	289.679	5,66	421.286	7,35	45,4
10 - 49	55.505	10,81	79.593	12,81	43,4	1.119.990	21,88	1.491.720	26,03	33,2
50 - 99	8.014	1,56	8.375	1,35	4,5	553.458	10,81	575.195	10,04	3,9
100 - 499	6.025	1,17	6.325	1,02	5,0	1.189.412	23,24	1.217.681	21,25	2,4
500 - 999	568	0,11	576	0,09	1,45	392.686	7,67	390.590	6,82	-0,5
1000 - altre	326	0,06	309	0,05	-5,2	826.243	16,14	746.182	13,02	-9,7

Fonte: Confindustria. Sesto Rapporto sull'industria italiana.
Roma, 1985. PP. 101-103.

sono passati da 188.667 a 252.975 e l'occupazione da 1.755.125 a 2.345.829. Il numero di addetti impiegati in queste unità ha raggiunto così il 40,9% delle forze di lavoro occupate nel settore manifatturiero.

Alla fine del periodo l'incidenza delle unità con uno o due addetti è diminuita, mentre è rimasta pressoché costante l'occupazione relativa. Pure le classi dimensionali superiori hanno registrato una perdita di peso in termini di quote relative. Non solo, ma le imprese con più di 500 addetti hanno accompagnato il fenomeno con una caduta netta dell'occupazione.

Gli anni settanta confermano così alcune tendenze che erano già emerse nei due decenni prima: il rafforzamento delle classi dimensionali intermedie (soprattutto di quelle con 10-49 addetti) e la progressiva diminuzione di importanza delle attività artigianali. Continua pertanto il processo di « travaso » delle unità piccolissime nelle unità medio-piccole e di converso il ridimensionamento del peso delle unità locali mediograndi e grandi.

Rispetto al periodo immediatamente precedente (1951-1961) si registra poi un certo rallentamento del tasso di crescita dell'occupazione industriale. Tanto che, mentre l'espansione delle unità locali segna un *trend* in continua ascesa, l'evoluzione degli addetti occupati nell'industria appare caratterizzata da un *trend* in netto declino.

Anche la dimensione media, per effetto di una molteplicità di fattori, si è notevolmente ridotta. L'aumento del numero degli stabilimenti, la generale tendenza all'abbandono delle grandi dimensioni, il progressivo processo di

TAB. 2 - Ampiezza media delle unità locali per classi di attività (1971-1981).

Rami di attività economica	1971	1981
Energia, gas acqua	92,9	108,5
Ind. estr. combust. solidi	645,8	275,9
Cokerie	382,4	404,4
Ind. Estr. petrolio e gas	318,1	128,4
Inc. petrolifere	206,0	305,5
Ind. sei combustibili nucleari	—	153,8
Ind. prod. dist. energia elet. gas acq.	93,8	104,8
Ind. raccolta depur. distr. acque	20,7	23,2
Ind. estr. manifatt. trasf. min. Chimiche	172,8	130,0
Ind. estr. prepar. min. metalliferi	245,5	178,4
Ind. prod. e prime trasf. metalli	681,4	589,8
Ind. estr. min. non metallif. non ener.	22,5	16,8
Prod. di prodotti in ceramica	49,1	43,4
Industrie chimiche	272,9	195,0
Ind. prod. fibre artificiali e sint.	1550,4	708,0
Ind. manif. lav. e trasf. metalli mec. prec.	137,8	99,3
Ind. della costr. di prod. in metallo	25,9	26,7
Ind. costr. install. macch. mat. mecc.	101,3	74,1
Ind. costr. install. rip. macch. uff.	841,4	346,5
Ind. costr. install., rip. moter. ettro	247,5	145,4
Ind. costr. mont., autoveicoli e acc.	2013,7	1465,7
Ind. costr. manut. autoveicoli e acc.	556,7	472,6
Ind. costr. strum. e appar. di precisione	75,3	35,9
Ind. manif. alim. tess. pelli legno	35,9	25,6
Ind. alimentari di base	25,3	23,9
Ind. zucchero bev. altri alim. tabac	79,4	94,3
Ind. tessili	69,6	36,0
Ind. delle pelli e del cuoio	21,3	16,3
Ind. calz. abbigl. biancheria	28,5	21,8
Ind. del legno del mobile di legno	9,9	9,0
Ind. della carta stampa editoria	60,7	46,5
Ind. gomma manufatti di materie plas.	116,8	58,7
Ind. manifatturiere diverse	26,1	14,6

Fonte: Confindustria, Sesto rapporto sull'industria italiana,
Roma 1983, PP. 101-103.

« deverticalizzazione » del sistema produttivo e infine l'andamento dell'occupazione non potevano non incidere in maniera determinante sull'ampiezza media delle unità locali. Cosicché mentre negli anni sessanta la dimensione media era andata via via aumentando, nel decennio successivo essa è andata riducendosi; è passata gradualmente infatti da 75,6 addetti per unità produttiva nel 1971 a 56,4 nel 1981. La diminuzione delle dimensioni medie delle unità locali ha interessato la maggior parte dei settori, a eccezione del ramo energetico e di quello alimentare non di base (Tab. 2).

I cambiamenti di struttura sembrano così confermare il processo di trasformazione dell'apparato industriale verso nuove forme di organizzazione dell'attività produttiva, caratterizzata dal frazionamento dei processi e dalla notevole diffusione dell'imprenditoria minore.

In anni in cui molti comparti hanno conosciuto fasi di rapido e prolungato sviluppo con relativo ampliamento dei mercati, la convenienza alla « divisione del lavoro » fra imprese e alla conseguente specializzazione produttiva, si è fatta sempre più forte. Tanto da ritenere che nel tessuto economico si siano andati diffondendo fenomeni di differenziazione intraindustriale che hanno interessato soprattutto le piccole e medie dimensioni aziendali.

Una prima spiegazione di questi fenomeni può essere ricercata nel tentativo da parte di molte medie e grandi imprese di attenuare gli eccessivi irrigidimenti dell'integrazione verticale decentrando alcune attività, prima organizzate al proprio interno, a imprese di piccole dimensioni.

Una seconda spiegazione riguarda più propriamente la natura del progresso tecnico attuale. Verrebbe meno cioè la tradizionale relazione positiva tra progresso tecnico da incorporare e elevati livelli di investimento e ciò sul presupposto che i nuovi processi produttivi richiederebbero relativamente minori quantità di capitale fisso. In questo modo le imprese raggiungerebbero la dimensione minima ottimale (ottimo tecnico) a un livello produttivo decisamente meno elevato di quello che necessitava in passato ².

Una terza spiegazione può essere individuata nel carattere discontinuo dell'andamento della domanda. La preferenza alla dimensione verrebbe posta a immediato confronto con una domanda sempre più mutevole, diversificata e personalizzata. Di qui la necessità di disporre di impianti, che possono essere disattivati con una certa flessibilità.

Una quarta spiegazione (ma che tiene conto anche delle prime tre) discende infine dalla possibilità da parte delle imprese di organizzare in

² Il prof. D. CANTARELLI (1979, p. 609) definisce questa evoluzione del progresso tecnico come processo di « miniaturizzazione tecnologica ». Si veda anche CANTARELLI (1985, p. 455-457).

sistema ³ i propri processi produttivi, lavorando per fasi distinte di produzione o specializzandosi per attività e attuando così una vera e propria divisione del lavoro fra imprese.

Si verrebbe così a creare una molteplicità di sistemi di piccole imprese integrate, capaci di raggiungere livelli ⁴ di efficienza e di competitività superiori alla media nazionale. Questa specializzazione sarebbe poi favorita sia dal flusso continuo di innovazioni che queste unità sarebbero in grado di incorporare ⁵, sia dalla possibilità di utilizzare una forza lavoro sempre più qualificata ⁶. E così mentre la specializzazione produttiva si rifletterebbe sulla « qualità del lavoro » impiegato dalle imprese, la migliore « qualité des hommes » amplificherebbe i risultati della specializzazione ⁷.

3. *Diffusione di imprese e decentramento produttivo nella Economia della produzione di Marshall*

Quando Marshall enunciò il principio ⁸, secondo cui molti vantaggi della divisione del lavoro possono essere conseguiti da piccole imprese a condizione che ve ne sia un numero elevato nella stessa attività, egli mise in luce perlomeno una prima fondamentale verità. E cioè che la dimensione in sé non avrebbe potuto discriminare a priori in ordine all'efficienza e che, ferme restando determinate condizioni, i vantaggi della grande dimensione

³ MOMIGLIANO (1980). Si veda anche FREY (1974).

⁴ Questi sistemi di imprese si troverebbero però a coesistere con un numero imprecisato di piccole imprese arretrate, sostenute da semplici economie di flessibilità e di evasione, caratterizzate da bassi indici di produttività.

⁵ In questo senso si veda BRUSCO (1975). Osserva infatti l'autore che non si può identificare la piccola impresa con l'arretratezza e come « Il decentramento produttivo serve ad affidare alle piccole imprese il controllo della maggiore quota possibile di forza lavoro senza per questo diminuirne la produttività ».

⁶ « En Italie la petite et moyenne entreprise grâce à sa plus grande flexibilité par rapport à l'évolution conjoncturelle réussit à échapper à ces contraintes du milieu qui privent la grande du consensus et de stabilité socio-politique nécessaires; en outre le travail spécialisé mis directement en cause tende à s'identifier avec l'entreprise et avec ses objectifs » (MANFREDINI, 1981, p. 57).

⁷ GISCARD D'ESTAING (1980) non esita ad affermare che « tous les aspects quantitatifs sont, certes, importants mais la qualité des hommes, de leurs décisions, de leur esprit de coopération, d'efficacité, d'invention marqueront le succès de l'entreprise. Ces qualités sont indispensables dans tous les secteurs et services de l'entreprise ».

⁸ Si tratta sostanzialmente della rielaborazione del noto teorema smithiano, secondo cui la divisione del lavoro è limitata dall'ampiezza del mercato.

sarebbero stati conseguibili anche da un insieme di piccole imprese operanti nello stesso settore ⁹.

Il ragionamento di Marshall può essere così riassunto: dato un distretto di espansione di un certo comparto, al verificarsi di un aumento della domanda, potrà essere conveniente scindere le diverse fasi in cui si articola il processo produttivo, in diverse piccole imprese, ottenendo in questo modo il massimo di economicità.

La portata dell'enunciato è di particolare rilevanza almeno in due direzioni. Innanzitutto in ordine alla possibilità da parte delle grandi imprese di incorporare fasi produttive, che prima detenevano al loro interno, per conseguire economie di costo. In secondo luogo alla possibilità che entrino nel mercato nuove unità specializzate nella fornitura di componenti o semilavorati, che verranno acquistati da altre imprese del settore per la realizzazione del prodotto finito.

A Marshall a dir vero non era neppure sfuggito il fatto che queste imprese sarebbero state in grado di garantire l'avanzamento tecnologico della propria produzione e dei propri processi. Si legge infatti nell'*Economics of Industry* che un piccolo stabilimento che si dedica a una sola breve fase del processo produttivo può avere il macchinario migliore e più altamente specializzato ¹⁰.

Ma a ben guardare l'innovazione tecnologica rappresenta per queste unità un qualcosa di più. Una specie di *conditio sine qua non* che ne assicura la sopravvivenza. Dal momento che si troverebbero sempre esposte al rischio di essere superate ed eliminate dal mercato a causa della possibilità da parte dei loro clienti di creare un reparto integrato ¹¹.

L'idea di Marshall era però che le piccole fabbriche avrebbero potuto effettivamente competere con le grandi nella misura in cui le piccole fossero state « addensate in gran numero nello stesso distretto » ¹².

Cosicché potrebbe sembrare che la localizzazione dell'attività produttiva rappresenti per le minori dimensioni come una sorta di ulteriore requisito per conseguire i vantaggi di una produttività crescente. In realtà, come giustamente è stato fatto notare, « l'incremento di produttività non è correlato funzionalmente ad un avvicinamento o allontanamento spaziale di imprese, ma, ad una crescente suddivisione interna, per stadi, dell'industria » ¹³.

Nel pensiero di Marshall cioè la maggiore localizzazione dell'industria

⁹ MARSHALL e PALEY MARSHALL (1975, pp. 73 ss.).

¹⁰ Ibidem, p. 74.

¹¹ SALLEZ (1979, pp. 134-135).

¹² MARSHALL e PALEY MARSHALL (1975, p. 73).

¹³ BECATTINI (1975, p. 10).

si pone come conseguenza logica dell'aumento di produzione. E poiché le piccole imprese nel nostro caso costituiscono a livello tecnico fasi distinte di uno stesso processo produttivo, sarà molto probabile che esse insistano su una stessa area di diffusione ¹⁴.

Ma il teorema presenta altresì un ulteriore interessante corollario. Scrive infatti Marshall: « se vi sono molte fabbriche, grandi e piccole, tutte impegnate nello stesso processo produttivo, sorgeranno industrie ausiliarie per soddisfare i loro « bisogni particolari » ¹⁵.

Come dire che, arrivati a un certo punto nello sviluppo di un settore, le imprese troveranno conveniente non solo scorporare fasi del proprio processo produttivo (sempre nella misura in cui questo sia possibile), ma addirittura affidare all'esterno attività ausiliarie che prima risultavano integrate con la funzione fondamentale. Supposto infatti che l'impresa sia definita come una unità impegnata nell'organizzazione di una serie di attività (acquisto di materie prime, trasformazione delle materie prime in semilavorati e dei semilavorati in prodotti finiti, vendite, pubblicità, trasporto e ricerca), ne deriverà che essa potrà specializzarsi in una o più di queste attività e acquistare direttamente dal mercato i prodotti e i servizi che prima produceva al suo interno.

4. *La motivazione tecnologica*

La strategia di autonomizzazione delle diverse funzioni del ciclo produttivo può porsi, come si è visto più sopra, sia nell'ambito di una politica di decentramento effettuato dalle grandi imprese verso le piccole sia viceversa nell'ambito di una politica di riorganizzazione dell'attività produttiva effettuata esclusivamente all'interno delle piccole e medie imprese.

È vero che nel primo caso può trattarsi di operazioni di mero scorporo — senza capacità propulsiva — effettuate al fine di ovviare alle diseconomie che la grande dimensione comporta in termini di conflittualità sindacale e di manovrabilità della forza lavoro ¹⁶. Ma è vero anche che nella maggior parte

¹⁴ Trattandosi cioè di industrie ausiliarie e complementari, che per definizione devono operare in continua cooperazione, ci si dovrà attendere che esse siano fortemente concentrate spazialmente. Tuttavia sembra opportuno sottolineare che questa maggiore concentrazione riguardi soprattutto la cosiddetta « subfornitura di capacità » piuttosto che quella di « specialità ». Dal momento che in questo secondo caso, essendo, in genere, i rapporti fra i *partenaires* più regolari, anche le distanze fra le imprese interessate potranno essere decisamente maggiori (in questo senso si veda SALLEZ, 1979, p. 135).

¹⁵ MARSHALL e PALEY MARSHALL (1975, p. 73).

¹⁶ BERNARDI (1976, pp. 471-72).

dei casi le imprese troveranno economicamente vantaggioso separare i vari stadi del processo di fabbricazione o le varie attività essendo le funzioni stesse efficienti a scale operative di diverse capacità.

Il fondamento teorico di tale comportamento si trova cioè nel diverso andamento delle curve dei costi medi di ciascuna attività o fase di produzione.

Supponiamo¹⁷ infatti che una impresa organizzi al suo interno tre funzioni A, B, C, e che esse presentino rispettivamente i seguenti andamenti di costi medi: sempre decrescente (A), sempre crescente (B), tradizionale andamento ad U (C).

Nell'ipotesi considerata la scala efficiente minima dell'impresa può definirsi come quel livello di *output* che permette di minimizzare i costi unitari globali intesi come somma dei costi unitari delle varie funzioni al punto di produzione considerato. Il che equivale a dire che a quel livello di *output*, mentre alcune funzioni avranno raggiunto, per così dire, l'« ottimo tecnico », altre non lo avranno raggiunto affatto. Mentre cioè certe attività avranno conseguito tutti i vantaggi dei rendimenti crescenti altre non potranno beneficiarne, perché magari la scala produttiva al punto considerato è troppo limitata.

Ma, con l'espandersi dell'industria, può essere conveniente affidare all'esterno la funzione soggetta a rendimenti crescenti (A) in maniera tale che un'altra impresa si specializzi in essa. Quest'ultima produrrà anche per altre imprese del settore. Infatti il prezzo praticato dall'unità subfornitrice sarà nettamente inferiore al costo che l'impresa committente dovrebbe sostenere se integrasse la fase o l'attività.

Anche la funzione caratterizzata da costi crescenti (B) può essere abbandonata o perlomeno ridotta nella sua scala di attività. L'impresa si rifornirà così da imprese più piccole e indipendenti oppure costituirà al suo interno impianti di piccole dimensioni.

Ora la validità generale di questo teorema è ancorata alla legge che lega l'evoluzione della specializzazione produttiva alla dimensione del mercato. Cosicché in prima approssimazione potrebbe sembrare che la disintegrazione verticale possa essere teorizzata solo in presenza di una industria in espansione. Ma è agevole ritenere che una impostazione di questo genere si riferisca fondamentalmente al caso in cui « il progresso tecnico si traduce esclusivamente in processi che impiegano una maggiore quantità di fattori fondo, ma non considerano il caso in cui il progresso tecnico assume caratteri diversi »¹⁸.

E cioè situazioni in cui si assiste a innovazioni nei processi produttivi

¹⁷ STIGLER (1951).

¹⁸ FORNENGO (1981, p. 231).

con relativa riduzione dell'impiego di materiali o di vera e propria eliminazione di intere fasi del ciclo di produzione; dove si intensificano le tecnologie relative all'applicazione dell'elettronica e quelle relative all'utilizzo di nuove fonti di energia; dove si applicano nuove tecniche di progettazione e di produzione. Perché in questi casi è fondato supporre che, indipendentemente dal fatto che il mercato si espanda, l'impresa sia egualmente spinta a decentrare le proprie attività e ciò nella misura in cui ricaverà economie di costo. Questa ipotesi interpretativa è inoltre suffragata dalle analisi empiriche condotte sull'argomento. Dato che sembra ormai acquisito — sul piano pragmatico oltre che su quello teorico — che la natura del progresso tecnico attuale sia tale, almeno in molti casi, da richiedere minori quantità di investimento (¹⁹).

In conclusione, se l'unica motivazione che guida l'imprenditore è quella tecnologica, è di tutta evidenza che per l'impresa sarebbe più conveniente specializzarsi in uno o più di queste attività e incorporare quelle che sono soggette a rendimenti crescenti o che invece registrano diseconomie di scala. Si assisterebbe così ad un primo tipo di decentramento produttivo caratterizzato da una disintegrazione verticale delle funzioni con relativa specializzazione delle imprese per attività.

Ma se il ciclo di produzione è segmentabile in più stadi aventi diverse scale ottimali minime, potrebbe essere conveniente per l'impresa dedicarsi a una o più di queste fasi. Tanto che prenderebbe vita una seconda forma di disintegrazione verticale non più per funzioni, ma per fasi produttive. In questo caso il calcolo delle economie di scala non verrà più effettuato in relazione all'intero processo produttivo, bensì in funzione dello stadio ottimale di ciascuna fase del processo che è possibile decentrare.

5. *Note conclusive*

Fino a questo punto la nostra analisi ha cercato di evidenziare come esistano diverse tendenze al raggiungimento delle dimensioni « ottimali » e ciò in dipendenza di una molteplicità di fattori, come l'appartenenza dell'impresa a un dato settore, lo sviluppo del settore stesso e il tipo di tecnologia che è possibile incorporare. E che queste diverse tendenze possono muoversi verso l'alto, cioè verso imprese di grandi dimensioni (integrazione verticale)

¹⁹ Esiste ormai una certa evidenza empirica che mostra la notevole capacità di diffusione orizzontale di molte nuove tecnologie (informatica, telematica, robotica microelettronica etc.). Si tratta cioè di innovazioni che presentano i caratteri di « *capital saving* » che abbassano, anziché rialzare, i minimi quanti.

oppure verso il basso (decentramento produttivo). In maniera tale che il « movimento » delle imprese verso l'ottimalità non si configura come un « movimento » unidirezionale, del tutto scontato, verso la grande dimensione, bensì come un processo di continuo adattamento delle loro strutture organizzative rispetto alle esigenze del sistema economico nel suo complesso ²⁰.

Va da sé che la dimensione, intesa in senso dinamico, si caratterizza allora come una sorta di variabile dipendente, soggetta, non solo alle propagazioni del sistema, ma anche al processo di accumulazione delle risorse (materiali e personali) che una certa impresa ha saputo realizzare in un dato lasso di tempo. Pare a noi cioè che la capacità della singola unità produttiva di affrontare i costringimenti dell'ambiente dipenda in primo luogo dalle potenzialità già insite nella sua struttura produttiva. Cosicché il fatto di conservare nel tempo una « dimensione ottimale » risulta funzione innanzitutto della quantità di risorse di cui l'impresa può già disporre.

Un tale modo di procedere ci offre la possibilità di individuare nella dimensione una variabile in continuo mutamento, che ammette uno sviluppo positivo, nullo o negativo ²¹; che accanto all'aumento della propria struttura tecnico-organizzativa, ammette pure il mantenimento di quella struttura oltre che la sua diminuzione.

Di conseguenza la dimensione aziendale non risulta più legata alle limitazioni della teoria tradizionale secondo cui, se si ipotizza una funzione di produzione omogenea, la dimensione dell'impresa è indeterminata. In realtà le variabili esogene (mercato, tecnologia, assetto istituzionale etc.) agiscono rispetto alla dimensione aziendale come forze centrifughe o centripete.

Tanto che se prevalgono le prime l'impresa sarà spinta a decentrare le attività o le fasi di produzione (e quindi a diminuire la propria struttura organizzativa), mentre se prevalgono le seconde l'impresa troverà economicamente vantaggioso integrare quelle stesse fasi o attività. Cosicché, se in un determinato momento risulta conveniente all'impresa specializzarsi per fasi produttive e per attività, essa si integrerà in sistema con altre imprese del settore. Fra queste imprese si instaureranno rapporti di produzione, che in realtà si concretizzeranno in rapporti tra sistemi parziali di un unico insieme organico complesso. Tra i sistemi parziali intercorrerà poi un'azione recipro-

²⁰ Come sottolinea infatti la Prof. Manfredini (1969, p. 152) « Le azioni creatrici delle micro-unità risalgono alle decisioni della disposizione privata. ... La disposizione nella sua decisione è condizionata dall'ambiente, anche quando le sue azioni determinano in definitiva il mutamento dei dati di partenza ».

²¹ GOLFETTO (1980).

ca così stretta che spesso è difficile delimitarne le rispettive funzioni. Nel senso che ciascuna di queste funzioni non è concepibile nella sua normalità al di fuori di tutte le altre.

Se è vero cioè che sia i costi di organizzazione delle attività all'interno dell'impresa sia i costi d'uso del mercato non sono nulli, ne deriverà che per l'impresa esiste un punto oltre il quale integrare una ulteriore attività o fase produttiva significa sostenere costi maggiori rispetto a quelli che si avrebbero se si facesse uso diretto del mercato. In una parola, per l'impresa si verrebbe a individuare una posizione, nell'ambito della propria organizzazione interna, in cui i vantaggi del *buy* sarebbero di tutta evidenza su quelli del *make*.

RIFERIMENTI BIBLIOGRAFICI

- BECATTINI G., "Introduzione" a Marshall A. e Paley Marshall M., *Economia della produzione*, Milano: Isedi, 1975.
- BERNARDI G., "Gli investimenti, fissi lordi industriali nel periodo del ristagno (1964-75)", *Rivista di Politica Economica*, aprile 1976.
- BRUSCO S., "Le condizioni di lavoro nelle fabbriche piccole, medie e grandi", *Inchiesta*, gennaio-marzo 1975.
- CANTARELLI D., "Dimensione delle imprese, progresso tecnologico e produttività dei fattori", *Il Risparmio*, 5/1979.
- , *Lezioni di economia politica*, Padova: Cedam, 1985.
- FORNENGO G., "I rapporti tra imprese", *Economia e Politica Industriale*, n. 30, 1981.
- FREY L., "Le piccole e medie imprese industriali di fronte al mercato del lavoro in Italia", *Inchiesta*, aprile-giugno 1974.
- GISCARD D'ESTAING O., "Les coûts comparatifs potentiels et le commerce international", sujet présenté au colloque international "Les formes actuelles de la concurrence dans les échanges internationaux", Ismea, Paris, 10-14 mars, 1980.
- GOLFETTO F., "La teoria dell'impresa esiste ancora", *Economia e Politica Industriale*, n. 28, 1980.
- MACHLUP F., "Teorie dell'impresa: marginalistiche, comportamentistiche, manageriali", in G. Zanetti (a cura di), *Contributi per un'analisi economica dell'impresa*, Napoli: Liguori, 1980.
- MANFREDINI M., *Il capitale nella disposizione pubblica e privata*, Padova: Cedam, 1969.
- , "L'importance prépondérante de la qualité du travail dans le commerce international et l'avantage comparatif potentiel", *Rivista Internazionale di Scienze Economiche e Commerciali*, gennaio-febbraio 1981, 28, 52-60.

- MARSHALL A. e PALEY MARSHALL M., *Economia della produzione*, Milano: Isedi, 1975.
- MARTELLARO J.A., "The Survival of Small Business Revisited", *Rivista Internazionale di Scienze Economiche e Commerciali*, giugno 1982, 29, 552-73.
- MOMIGLIANO F., "Piccola impresa come sistema", *Economia e Politica Industriale*, n. 27, 1980.
- SALLEZ A., "Subforniture, produttività del sistema industriale e sviluppo economico regionale, in R. Varaldo (a cura di), *Ristrutturazioni industriali e rapporti fra imprese*, Milano: Franco Angeli, 1979.
- STIGLER G.J., "The Division of Labor is Limited by the Extent of the Market", *Journal of Political Economy*, June 1951.

SIZE IN THE THEORY OF THE FIRM

The object of this paper is the analysis of size in the theory of the firm. Its main points are:

- size itself cannot completely explain the degree of efficiency of a production unit;
- there are different paths to achieve optimal size;
- Marshall's theory (according to which many advantages of the division of labour can be achieved also by small factories which are assumed to be numerous in every industry) forms the nucleus of a more general theory of business relationships among firms.

I GRANDI SAGGI DI GIUSEPPE DI NARDI *

di

ACHILLE AGNATI **

Questa raccolta di saggi – pubblicati dal 1933 al 1983 – propone temi teorici e questioni pratiche dell'economia politica in un complesso monumentale di quasi 1500 pagine. Il lettore, già ad apertura di libro, può indugiare sulla pregnanza delle suddivisioni editoriali (da noi virgolettate) e dei titoli degli scritti secondo raccomanda l'«Avvertenza» di Carlo Pace invitando a chiarire e a calibrare, ma anche a definire, la qualità di prim'ordine – nelle sue novità e nelle sue sottigliezze – del pensiero economico di Giuseppe Di Nardi.

Il 1° volume *La revisione metodologica* esordisce con «La contestazione» facendo perno su *La svolta degli anni Trenta: introduzione generale* (1981) che – essendo la lezione di congedo dall'Università – si presenta come bilancio e testamento scientifico onde converrà dirne alla fine di questa lunga recensione volutamente espositiva, piuttosto che critica, trattandosi di un corpus magistrale in sintonia – nonostante particolarità personali – con la scuola di Demaria. Recensione espositiva condotta attraverso saggi chiave da noi trascelti nella quasi totalità della maggiore saggistica dell'A. qui ristampata, ma che il lettore interessato potrà quando sostituire quando arricchire con altri saggi richiamati dal proprio visus culturale.

Ne «L'ordine come organizzazione funzionale dell'economia», Di Nardi con *Il relativismo della scienza economica* (1948) conferma le idee fondamentali di chi – cresciuto nel clima d'opinione degli studi sull'equilibrio generale, ma sostanzialmente nutrito del seme marshalliano – concepisce l'economia politica basata sull'unico principio unificatore dell'azione economica: il minimo mezzo che diventa ricerca del massimo di benessere onde il significato delle leggi economiche è la conoscenza dei meccanismi empirici della vita economica. Sui due principi fondamentali, rarità delle risorse e valore pratico della conoscenza, Di Nardi costruisce la sua posizione metodologica. Posizione critica perché riconosce subito che il principio del minimo mezzo come criterio interpretativo dell'azione economica dell'uomo, e dunque quale criterio esplicativo della realtà economica, comporta insufficienze onde il riconoscimento del principio d'indeterminazione introdotto in economia politica dal Maestro Demaria: «Da questo punto di vista i nessi causali non appaiono più certi, ma probabili; non appaiono più necessari ed

(*) G. DI NARDI, *L'economia a una svolta critica*, vol. I, *La revisione metodologica*, Milano, Giuffrè, 1984, pp. VIII-396. Lire 28.000. ID., *L'economia a una svolta critica*, vol. II, *Contrattazione collettiva e distribuzione dei redditi*. Ibidem, 1985, pp. VIII-340. Lire 27.000. ID., *L'economia a una svolta critica*, vol. III, *Moneta, ciclo, occupazione*, Ibidem, 1985, pp. VIII-364. Lire 28.000. ID., *L'economia a una svolta critica*, vol. IV, *Il dualismo nello sviluppo*. Ibidem, 1985, pp. XII-360. Lire 27.000.

(**) Università di Padova.

universali, ma contingenti e legati alla permanenza di certe condizioni che sfuggono alla sistematica ripetizione » (p. 58). Dalla rottura metodologica Di Nardi evolve gradualmente augurandosi che, in avvenire, venga alla luce un nuovo principio d'interpretazione che superi i due criteri irrinconciliabili del comportamento razionale e dell'indeterminazione dinamica.

E Di Nardi è diventato, infatti, via via meno pessimista sulle possibilità di progresso della scienza economica: ne *L'economia politica, scienza del benessere collettivo* (1954) si danno cinque indirizzi principali del pensiero economico moderno, dorsale di un'economia passata – nell'arco temporale che Di Nardi ha vissuto scientificamente – da scienza astratta del comportamento umano a scienza pratica del benessere collettivo: la « nuova economia del benessere »; la « programmazione economica »; l'« economicità »; la « teoria macroeconomica dell'occupazione »; la « scienza della contabilità nazionale ». Di Nardi è, in linea di principio, favorevole a questi cinque indirizzi e – pur riconoscendo che la scienza economica tende a diventare ermetica – ritiene questo ermetismo frutto d'un maggiore approfondimento che consente un accresciuto potere d'intervento e d'azione all'uomo di governo. Onde, su questa strada, ampia pur se argomentativamente e analiticamente composita, vanno letti i saggi *L'analisi funzionale per la comparazione dei sistemi economici* (1964); *Le funzioni di un sistema dei prezzi* (1965); *L'economia come organizzazione funzionale* (1967).

« L'indeterminazione » con *Interdipendenza e indeterminazione dinamica nella teoria economica* (1942) dice la doppia impronta della teoria dell'equilibrio e del principio d'indeterminazione secondo legami che uniscono la teoria d'interdipendenza generale dei fenomeni al principio dell'equilibrio economico, giacché il principio d'interdipendenza delle grandezze economiche s'oppone al principio di causalità; la teoria paretiana, che riunisce il principio d'interdipendenza e il principio d'equilibrio, contiene in sé stessa una contraddizione implicita. L'errore della concezione paretiana – è qui il nocciolo della concezione di Demaria – sta nell'avere inteso il tempo come « vuoto di fatti nuovi » che disturbano le posizioni di partenza e conducono il sistema a posizioni sempre più lontane dalla posizione d'equilibrio ideale. Di Nardi si augura che il tipo di conoscenza cui è impari l'analisi deduttiva possa essere agevolato dalla raccolta di informazioni induttive: « Per questo lo studio di particolari biografie di eventi, di molte biografie, malgrado le manchevolezze tecniche che ancora limitano le facoltà esplorative di questo metodo, può ridare all'indagine economica quel vivo, umano senso di concretezza che si era smarrito nel vacuo esercizio di costruire equazioni in numero pari alle incognite » (p. 183). In perfetta sintonia con questo saggio magistrale, si ha lo svolgimento su *Le istanze innovatrici della dinamica economica* (1953) dove « In definitiva, si tratta di riportare l'irrazionale nel razionale, di foggare cioè schemi mentali meno angusti, che è poi l'eterno problema della scienza sorta non per appagare la sete inestinguibile della conoscenza assoluta, ma per lo scopo più limitato e pratico di aiutare l'uomo a vivere la sua vita terrena, dominando come può le forze del creato e le passioni suscitate dai suoi istinti » (p. 198). Istanze ampiamente svolte in *Indeterminazione ed esogeneità* (1979) dove compare già articolato il discorso categorico di Demaria sulle variabili extraeconomiche.

« Lo storicismo dell'economia » affronta *Fatti e teorie nello svolgimento del pensiero economico* (1938) specificando, molti anni dopo tematicamente e sotto specie analitica, *Dalla scienza della ricchezza alla teoria delle scelte* (1959): in questa prospettiva, l'economia teorica dovrà rinunciare a una costruzione edificata su un principio unico riconoscendo che le diverse parti del sapere partono da una regola d'interpretazione diversa caso per caso. È il principio delle spiegazioni complementari, preso a prestito dalle scienze naturali, che in Di Nardi ha due fonti: la sistematica di Pareto dell'equilibrio economico generale accettata come schema teorico generale in cui si inquadra ogni fenomeno concreto; la sistematica di Demaria sull'indeterminazione secondo cui nel tempo reale, che è storico ed economico insieme, non v'è mai certezza di

prevedibilità e di predeterminazione dei rapporti economici. E l'originalità di Di Nardi sta nel tentativo di conciliare la concezione rigidamente meccanicistica della teoria dell'equilibrio economico generale, propria di Pareto, alla concezione di un mondo economico in perpetua evoluzione, propria di Demaria. Di Nardi cerca le ipotesi di base sulle quali lo schema logico paretiano raggiunge i risultati di determinazione, di stabilità, di equilibrio; rileva le ipotesi limitatrici del sistema dell'equilibrio generale; in breve, getta i fondamenti d'un rinnovamento della teoria economica sempre ricordando che — laddove il perseguimento di un obiettivo si realizza senza essere condizionato, walrasianamente, dalla realtà delle risorse — la scelta scompare e ci si trova al di fuori dell'economia, giacché si è nel campo o della tecnologia o della precettistica: « L'atto della valutazione segna dunque il momento differenziatore dell'attività economica nel più vasto campo dell'azione umana » (pp. 243-44). E in questa luce, logicamente già plurimodale, rientra la varietà degli scritti *Economia e senso comune* (1970); *Antilogie dell'economia contemporanea* (1973); *È obsoleta la scienza dell'economia?* (1974); *Una riflessione sul pensiero smithiano* (1976); *Keynes e l'analisi di qualità* (1976); *Pantaleoni rivisitato* (1977); *Sviluppo industriale e qualità della vita* (1977); *L'economia fra tradizione e innovazione* (1982); *Una postilla alla concezione dello stato fattore di produzione* (1983). Tutti saggi coerentemente cresciuti sulle linee precedenti nel quadro totalizzante che ricomprende autorità quali Smith, Pantaleoni, Keynes; temi perenni d'incivilimento come la qualità della vita; argomenti categorici nuovi come lo stato fattore.

Il 2° volume *Contrattazione collettiva e distribuzione dei redditi* prova — attraverso problemi pratici — lo sbocco del processo economico nell'indeterminazione della quale Di Nardi ha affrontato un caso tipico col monopolio bilaterale — sua tesi di laurea nel 1932-33 — percorrendo l'analisi economica del soggetto secondo il calcolo di convenienza individuale, rifacendosi a Edgeworth a Marshall a Jannaccone, ma anche e fondamentalmente alla teoria del prezzo di Walras. La convinzione di Di Nardi è che il prezzo di equilibrio mai si raggiunge essendo la sua immagine data dalla teoria del mercato di concorrenza perfetta; tuttavia, passando all'analisi dei mercati imperfetti, configurazioni più realistiche, egli esprime un giudizio di fattibilità sul funzionamento dell'economia empirica sempre anomala rispetto allo schema logico del mercato perfetto. Ecco quindi raccolte le tappe del cammino di Di Nardi nell'analisi delle forme di mercato e delle loro implicazioni in sede distributiva, onde dalla successione degli scritti si può trarre la configurazione di un sistema economico che, non avendo la rigidità del meccanismo deterministico, è condizionato da vincoli che riescono a tenere insieme un organismo. Tale è pure il significato dell'equilibrio economico inteso quale rispetto delle « compatibilità » fra le istanze contrapposte di contraenti mossi dal proprio interesse, il che porta ad ammettere l'esistenza di più valori contemporaneamente possibili come risultato dello scambio onde una fascia di valori compresi tra un minimo e un massimo, ossia una zona di oscillazione dei valori di equilibrio. Questa la concezione di equilibrio organico — sostenuta nei saggi raccolti in questo 2° volume — contrapposta alla concezione di equilibrio meccanico.

Secondo Di Nardi — ed è un po' il leit-motiv ideale e analitico del suo pensiero economico — schema teorico e realtà pratica si distaccano in tre punti fondamentali: la distribuzione iniziale dei redditi, giacché ogni schema di concorrenza perfetta restà razionale qualunque la distribuzione dei redditi di partenza onde lo schema è formale non implicando alcun giudizio di valore sulla ripartizione del benessere fra i diversi componenti la collettività; gli scopi del fenomeno economico, giacché mentre l'azione individuale è destinata a soddisfare un interesse personale, gli scopi umanitari, sociali, politici possono richiedere per il loro conseguimento l'intervento dello stato e allora l'ottimo non è più quello dello schema di concorrenza perfetta; la forma del mercato che, allontanandosi dall'ideale della concorrenza perfetta, consente allo schema teorico di segnalare ogni possibilità di inefficacia e di spreco.

Di qui le ragioni teoriche, analitiche, metodologiche degli « Schemi di riferimento » con il già citato *La indeterminazione nel monopolio bilaterale* (1934). Di Nardi accetta la conclusione che le soluzioni di Marshall e di Edgeworth portano a una soluzione di indeterminazione, ma rileva che l'assenza di determinazione dipende dall'ipotesi implicita secondo cui due contraenti detengono la medesima forza contrattuale perché — altrimenti — uno dei due finirebbe per imporre all'altro il proprio principio d'utilità. Il risultato ultimo è che l'analisi del monopolio bilaterale si fonda così sull'analisi del monopolio unilaterale: il monopolista debole finisce per confondersi con il consumatore che subisce il prezzo del monopolio unilaterale. Accettata la possibilità di paragonare i benefici in termini di utilità di diversi individui, Di Nardi vede la misura della forza contrattuale nel rapporto tra beneficio tratto dallo scambio e quantità scambiate. L'analisi completa dello scambio può tuttavia essere condotta soltanto conoscendo la forza contrattuale delle due parti; pure — osserva Di Nardi — si può pensare che, in moltissimi casi, le forze contrattuali delle due parti siano interdipendenti ad es. nei contratti tra lavoratori e datori: l'aumento della forza contrattuale dei datori spinge i salari al limite minimo influenzando sulla forza contrattuale futura dei lavoratori. Quando l'analisi viene estesa alla teoria dinamica del monopolio bilaterale, la possibilità di scoprire una posizione d'equilibrio determinato scompare: onde la soluzione del problema non ha carattere di certezza ma solo di probabilità. Su questa linea di pensiero procedono gli scritti: *Il salario. Teoria e realtà concreta* (1935) e la corposa recensione critica *La teoria neoclassica della distribuzione nelle « Lezioni » di C. Bresciani Turroni* (1949) entrambi saggi tra loro ben diversi, ma pertinenti alla tematica argomentativa e dialettica della parte del volume.

Saggi cui si legano quelli che, in sede pratica e di logica del movimento economico, affrontano specificazioni sempre problematiche sulla « Dinamica dei salari e politica dei redditi » con *I limiti di oscillazione del salario nei rapporti collettivi di lavoro* (1938) e, successivamente, *Sviluppo economico e distribuzione del reddito* (1957); *Gli effetti del progresso tecnico sulla distribuzione del reddito* (1961); *Lineamenti generali di « politica dei redditi »* (1965); *La politica dei redditi in Italia* (1966); *Partecipazione e autogestione* (1976); *L'impresa pubblica manifatturiera* (1977); *L'impresa fra socialità ed efficienza* (1977).

Il 3° volume *Moneta, ciclo, occupazione* dice come in Di Nardi la teoria dei fenomeni monetari si sia sempre bagnata nella loro storia: le vicende monetarie degli anni Trenta sono annodate ai tentativi di Inghilterra, Stati Uniti, Francia per affermare in proprio il controllo dei mercati finanziari non portando — come desiderato — all'unificazione del sistema dei pagamenti internazionali con una moneta di conto accettata per le sue garanzie di stabilità. Si sono avuti esperimenti di varie forme di « scambi compensati », ossia di chiusure dei mercati valutari e di controllo dei movimenti di persone, capitali, merci, con generale restrizione degli sbocchi e cospicua riduzione dei tassi di occupazione secondo il verbo economicistico tratto dalla nuova ortodossia della *Teoria generale dell'occupazione, dell'interesse e della moneta* di J.M. Keynes.

Tre saggi di riferimento sono fondamentali nel riconoscimento post data dello stesso Di Nardi. La decisa opposizione degli Stati Uniti di Roosevelt alla politica dell'Inghilterra di svalutazione dell'argento — causa le ripercussioni negative per i rapporti con la Cina e con i produttori di argento negli Stati Uniti — è analiticamente argomentata ne *Il problema dell'argento e i provvedimenti del Presidente Roosevelt* (1934); la moneta-lavoro escogitata in tempo di guerra per far credito al potenziale economico della Germania — che difendeva il potere di acquisto della carta moneta nazionale con pesanti controlli al movimento delle merci e con patti di vantaggio coercitivamente unilaterale per l'approvvigionamento di derrate alimentari e materie prime tramite compensazioni bilaterali, con gli stati dell'Europa Orientale, smerciando le eccedenze di beni non necessari — è in *Oro e lavoro nella nuova economia* (1941); infine, ancora dell'opposizione statunitense alla politica inglese di svalutazione dell'argento e della reazione degli ambienti finanziari europei alle proposte di Cassel alla Conferenza monetaria di

Genova del 1922 – per la determinazione del cambio mediante il rapporto dei poteri di acquisto delle monete inconvertibili – si legge in *Sulla indeterminazione « a priori » delle parità dei cambi* (1948).

Sovente dalle politiche economiche del breve andare nascono teorie temporalmente e argomentativamente limitate perché valgono finché non sorgono reazioni empiriche distruttive degli assetti cui mirano quelle politiche economiche che oggi si dicono di periodo lungo. Queste le teorie limitate o parzialmente inaccettabili delle situazioni di equilibrio, le quali – per i fenomeni monetari – non sussistono perché, appena interviene il potenziale di credito generato dal sistema bancario (quale effetto del moltiplicatore dei depositi), si ha l'ineluttabile conclusione che l'equilibrio monetario è sempre indeterminato per le ragioni teoriche di cui distesamente nel saggio *La svolta degli anni Trenta: introduzione generale* (1981), che riprenderemo alla fine di questo scritto, e nelle molteplici applicazioni nei saggi del già percorso 2° volume *Contrattazione collettiva e distribuzione dei redditi*. Ossia, sostanzialmente, perché il sistema economico reale, e per esso il connesso sistema economico teorico, fa che una rigorosa teoria generale al riguardo debba riconoscere il ruolo secondario delle variabili economiche endogene nei cambiamenti della dimensione e della causalità della vita economica con l'assoluta preminenza delle attività delle variabili esogene di propagazione nel determinare la dimensione e l'orientamento delle variabili economiche endogene. Sistema reale sempre immerso in un tempo reale mai vuoto di fatti nuovi che, come variabili esogene entelechiane, spiegano i processi violenti di sviluppo-inviluppo e di originalità della vita economica nonché gli inomogenei piani temporali della sua struttura monetaria (che, in termini di prezzi e loro numeri indice, è materia prima della metodologia denominata cinematica storica del movimento dei prezzi).

In termini generali, la storia dei fenomeni monetari – cui Di Nardi ha sempre guardato – può sembrare anomala, specialmente in Italia, per la presenza istituzionale del Ministero del Tesoro con la sua posizione sui generis fra le autorità monetarie cui è giuridicamente demandato il controllo sul sistema bancario; ma soprattutto può sembrare anomala per la caduta nello « stato assistenziale » supremo distruttore del risparmio privato attraverso trasferimenti per scopi desolatamente improduttivi. Negli anni Ottanta correnti anche le monete inconvertibili hanno consentito di contrastare l'inflazione e di mantenere una considerevole stabilità rispettando condizioni affidate in passato agli automatismi discendenti dalla convertibilità. Caduti gli automatismi, ritenuti incompatibili con l'obiettivo della piena occupazione delle risorse disponibili, la politica monetaria oggi giorno richiede più informazione, più vigilanza, forse più coercizione.

Di Nardi si preoccupa sempre di determinare in quale misura la politica degli Istituti sia stata conforme alla funzione razionale della Banca centrale; in quale misura questa politica abbia contribuito a sviluppare il reddito nazionale; in quale misura questa politica abbia invece ostacolato il progresso dell'economia italiana. Analizzando l'estensione della circolazione del credito, Di Nardi esamina i suoi rapporti con l'accrescimento del reddito reale e sottolinea come, nei periodi d'aumento del reddito, il credito si sviluppi nella sua quantità reale e non soltanto nella quantità nominale. Così – in questa precettistica e nella composita luce analitica sopra accennata – si presentano e lo spaccato dottrinario « Teorici della moneta e dei sistemi monetari » con *Nota sui concetti di velocità ed accelerazione della circolazione monetaria* (1934); e i già citati tre saggi: *Il problema dell'argento ed i provvedimenti del Presidente Roosevelt* (1934), *Oro e lavoro nella nuova economia* (1941), *Sulla indeterminazione « a priori » delle parità dei cambi* (1948); quindi anche *Saggio d'interesse, investimenti pubblici e investimenti privati* (1951); *I vincoli alla circolazione dei biglietti di banca nell'ordinamento italiano* (1954); *Il prezzo dell'oro* (1966); *La riforma del sistema monetario internazionale* (1973); *Relativismo e stabilità monetaria* (1975).

E ancora, nella stessa composita luce tematica e analitica, la spiegazione delle imprese bancarie in « Teoria e politica bancaria » con *La Banca d'Italia nel periodo 1894-1943* (1944); *L'intervento dello stato per la manovra del credito e degli investimenti* (1946); *Politica dello sconto ed elasticità della circolazione (Un caso storico)* 1954; *Trenta anni di « Bancaria »* (1977); *Se la moneta non contasse* (1982). Infine, la specificazione delle fluttuazioni e dell'occupazione in « Teorie del ciclo e politica dell'occupazione » con *Piena occupazione e scambi internazionali nell'Unione Europea* (1949); *Spese pubbliche e occupazione in Italia* (1953); *Investimenti in opere pubbliche per la stabilità dell'economia* (1954); *Risparmio e inflazione* (1975); *Ciclo e inflazione* (1974); *Schemi individualistici e correttivi keynesiani* (1979); *L'economia italiana nel biennio 1979-80* (1981); *La sfida al governo razionale dell'economia* (1982), tutti saggi che – pur di argomento monetario – chiariscono aspetti cardinali dell'economia reale italiana ed europea del secondo dopoguerra mondiale.

Il 4° volume *Il dualismo nello sviluppo* è sulle teorie dello sviluppo economico trattando particolarmente le esperienze italiane dell'intervento pubblico straordinario per correggere il ritardato sviluppo del Mezzogiorno. L'apparato teorico – filigrana analitica delle politiche per ridurre e annullare il dualismo nello sviluppo, ricorda Di Nardi – è il moltiplicatore di Kahn trasferito (in politica economica) dall'applicazione congiunturale, ossia di breve andare, al processo strutturale, ossia di lungo andare. Trasferimento impertinente perché l'incentivo alla domanda globale, aumentando la spesa pubblica, attenua e forse corregge lo sviluppo dualistico solo se il flusso del reddito aggiuntivo si consolida nella formazione di nuovi capitali; oppure se la disponibilità di nuove fonti di reddito si localizza nella regione a sviluppo ritardato.

L'esperienza di questo dopoguerra mondiale ha provato che tali condizioni in Italia non si sono verificate: quanto sia disatteso tale impegno politico economico si desume dai traguardi economici politici sociali puntualmente mancati e puntualmente rinviati. Pur non essendoci tutti gli scritti (sono elencati nella Bibliografia generale che chiude il 4° volume) sulle condizioni necessarie alla eliminazione del dualismo nello sviluppo, si intende come per Di Nardi il Mezzogiorno sia un caso particolare perché il Meridione è un insieme di regioni strettamente legate da vincoli politici, economici, amministrativi all'Italia settentrionale, territorio questo, d'intensa industrializzazione e in ulteriore espansione. Le spese pubbliche nel Mezzogiorno sono fonte di stabilità della domanda globale per tutto il sistema industriale italiano, e in particolare per quello delle regioni settentrionali che, disponendo di un'industria più sviluppata, traggono i maggiori benefici da un'espansione della domanda attraverso un meccanismo di moltiplicazione latamente inteso.

Ecco allora « Presupposti razionali e carenze operative della politica per il Mezzogiorno » con *La disoccupazione nel Mezzogiorno* (1951), mentre attorno al 1952 ancora si pensava che un intervento straordinario – sorta di una tantum nel settore dei lavori pubblici – fosse sufficiente a ristabilire la situazione nel Mezzogiorno giacché il problema si sarebbe dovuto affrontare seguendo una politica concertata d'investimenti coordinati. Onde *Precettistica intuitiva e valutazione econometrica per lo sviluppo delle aree depresse* (1952): nel corso degli anni successivi al 1950, Di Nardi ha approvato l'intervento dello stato nel settore delle infrastrutture, giudicandolo indispensabile per creare le prime condizioni favorevoli a una futura industrializzazione. E, quindi, tematicamente contigui di sostanza e di pathos: *Presupposti economici della trasformazione agraria nel Mezzogiorno* (1954); *Le politiche di impiego dei mezzi per lo sviluppo economico* (1956); *Sviluppo bilanciato e credito all'esportazione* (1957); *I provvedimenti per il Mezzogiorno: 1950-1960* (1960). Con *Esperienze italiane di sviluppo economico* (1963) dal punto di vista del Mezzogiorno si vede come i legami con le regioni settentrionali abbiano presentato due aspetti: l'economia industriale del nord, grazie al suo rapido sviluppo dopo le distruzioni della guerra, ha consentito la raccolta dei capitali necessari per investimenti massicci nel Mezzogiorno; d'altronde, un settore industriale in espansione, e teso ad aprirsi verso i mercati stranieri, ha

posto proprio alla politica di sviluppo del Mezzogiorno problemi particolarissimi che possono averne ritardato il completamento.

Sulla politica di sviluppo per il Mezzogiorno, Di Nardi parla in altro scritto di « ingenua trasposizione dalla teoria alla pratica »: un autentico piano di sviluppo per il Mezzogiorno non è mai stato progettato. L'aumento delle spese pubbliche, in sé stesso e per sé stesso, avrebbe potuto determinare un aumento della domanda globale; ma questa domanda, non sostenuta da una politica d'espansione generale e coerente, si sarebbe fatalmente e totalmente volta verso altri mercati. Si pensava invece che l'aumento della domanda avrebbe spinto le nuove unità di produzione a installarsi nell'Italia meridionale. Tuttavia, la teoria dello sviluppo economico è uscita arricchita, anche se non analiticamente irrobustita, da alcune esperienze della Cassa per il Mezzogiorno fra le quali spiccano: l'economia delle regioni settentrionali tesa ad attirare le risorse produttive delle altre regioni italiane; l'impotenza del programma d'infrastrutture pubbliche realizzato nell'Italia meridionale, sebbene accompagnato da impulsi all'industria privata, dinnanzi allo sviluppo dell'Italia settentrionale. L'unico elemento di speranza per il progresso delle regioni meridionali erano e forse sono le prospettive di nuovi mercati nei paesi mediterranei, purtroppo economicamente e politicamente via via meno affidabili. Ecco quindi le analisi e le speranze in *Leggi e modelli di sviluppo economico* (1956) e *Tempi brevi e tempi lunghi nelle politiche di sviluppo* (1983).

Sempre sulla stessa linea è così proficuo leggere — da « L'analisi economica dei progetti d'investimento » — « *Criteri* » e « *Indicatori* » per la scelta degli investimenti (1957); *Osservazioni sui criteri di scelta degli investimenti* (1964) e — da « Il dualismo nell'ambito della comunità internazionale » — *Le « zone marginali »* (1958); *I mezzi di intervento per la creazione di nuove attività* (1960); *L'agricoltura nella strategia dello sviluppo economico* (1962); *Lo sviluppo equilibrato delle regioni* (1967); *Il mito dello sviluppo generalizzato* (1977). Entrambi gruppi di saggi nei quali la maestria analitica si intreccia con pienezza di risultati conoscitivi all'esperienza dell'economista applicato.

Di Nardi antepone la lezione di congedo *La svolta degli anni Trenta* (1981) alla raccolta dei quattro volumi di saggi i quali costituiscono i riferimenti analitici ai giudizi di sintesi espressi in quello scritto introduttivo dove ci si riferisce alla revisione degli strumenti d'analisi dell'economia politica. Revisione cominciata all'inizio degli anni 1930 e non certo conclusa, mentre si ricorda che dal *Tableau Economique* (1758) di Quesnay al *Manuale* (1906) di Pareto una teoria generale dei prezzi di mercato è l'unica rappresentazione generale dell'esperienza umana tratta dall'attività economica. E — parafrasando appena alcuni passi della lezione di congedo — ricordiamo che Di Nardi identifica quella revisione metodologica, cui si iscrive la raccolta di saggi, con la svolta degli anni Trenta giacché il perfezionamento formale della teoria nella veste matematica dell'equilibrio economico generale ne ha accentuato l'anima deterministica, già apparsa nelle trattazioni dei classici, mediante sillogismi di tipo causale. Nello schema logico neoclassico, il sistema economico è un insieme di meccanismi legati da nessi di necessità onde si ammette il carattere naturalistico delle relazioni tra i fenomeni dell'economia, carattere che acquistando precisione e coerenza formale ha imposto tali semplificazioni del mondo reale da renderlo pressoché irricognoscibile. In breve, meccanismo e organicismo, determinismo e indeterminismo, certezza e probabilità delle sequenze quando logiche quando sperimentali, sono sia modi d'intendere l'attività economica sia categorie che non escludendosi o elidendosi si completano nell'economia politica la quale è scienza storica più che scienza naturalistica: rapporti materiali e rapporti immateriali che, riguardando la vita, sono di natura organica e biologica più che di natura meccanica. L'economia politica è, dunque, scienza storica più prossima alla biologia (Marshall) che alla fisica (Georgescu-Roegen): ogni accadimento dell'economia, come i fatti della biologia, sono la risultante dell'eredità e delle mutazioni; del passato e

di quel presente che risulta difforme dal preesistente condizionamento. Questo perché l'eredità favorisce il meccanismo nella sua esclusiva capacità di determinazione, ma i fatti nuovi ossia gli entelechiani, le mutazioni imprevedibili, rendono possibili sequenze quando logiche quando sperimentali diverse da quelle imputabili al meccanicismo. Perciò, in economia politica, il futuro non è predeterminabile dalle forze dello stato iniziale, che l'osservatore-studioso individua e determina come se potesse dividere il presente dal passato, il contingente dalla necessità. Il peso della tradizione suggerisce un ordine sistematico che regge l'economia politica come scienza, ma nel sistema economico empirico le mutazioni siano esse imprevedibili o prevedibili non distruggono l'ordine che l'economista è tenuto a rinvenire.

Questa la visione scientifica di Di Nardi che – conscio del realismo dell'economia politica – nota la compatibilità fra le grandezze del sistema reale e la zona d'indeterminazione (intesa come fascia di valori) di ciascuna variabile e quindi dei limiti di oscillazione della variabilità, limiti determinati dalla condizione che le mutazioni siano contenute in dimensioni che non distruggano la compatibilità, in cui essenzialmente consiste l'ordine. Fuori di questi limiti, si ha la rottura del sistema, che nel campo delle relazioni umane è la rivoluzione. Questa è la concezione dell'ordine che ha orientato gli studi di Di Nardi sin dalla tesi di dottorato del 1932-33 su *L'indeterminazione nel monopolio bilaterale* e dal saggio del 1938 su *I limiti di oscillazione del salario nei rapporti collettivi di lavoro*. Onde, a conclusione di questa lunga recensione espositiva, nella lezione di congedo *La svolta degli anni Trenta* (1981) si ritrovano i semi di cinquant'anni prima: « L'analisi dei limiti di oscillazione porta a schematizzare un tipo di ordine, concepito come organizzazione funzionale dell'economia. Ogni soggetto di decisioni economiche è un operatore che rende servizi ad altri soggetti, o alla comunità nel suo complesso e ne percepisce in contropartita una retribuzione. L'ordine permane fino a quando le retribuzioni sono tali da rendere attiva l'offerta dei vari servizi da una parte e la domanda degli stessi dalla parte opposta. Se i termini dello scambio restano alterati dalla presenza di forze egemoniche, il sistema degrada nel disordine, con l'aumento crescente dell'entropia » (p. 30). Precettistica teorica e conoscitiva d'indubbia validità nello spazio-tempo che stiamo vivendo.

Dunque, attraverso 70 scritti dal 1933 al 1983 si seguono storia e teoria condotti per mano da « Settanta Saggi » che raccolgono, in una grande saga, storia e teoria dell'economia politica degli ultimi cinquant'anni: una sorta di Bibbia per intendere corpo e specificazioni dell'economia teorica e dell'economia applicata. Di Nardi ha innestato la sua formidabile preparazione scientifica in una vasta conoscenza pratica tratta dalla partecipazione a organismi economici italiani ed europei. Di questa produzione scientifica denominatore comune è il realismo in una scienza economica tesa all'astrazione; strumentario analitico è il rigore metodologico contro la superficialità della sterminata pubblicistica economica e finanziaria. Nato alla scienza negli anni della « high theory » dell'economia pura, Di Nardi indica il « mare magnum » tra lo schema ideale e l'esperienza empirica: specialmente la sua ricostruzione della teoria dei prezzi e dei mercati rivaluta i postulati teorici d'una dottrina, in sede economica, troppo superficialmente adottata e, in sede politica, troppo disinvoltamente applicata.

In breve, la lezione professata da Di Nardi richiama sia la realtà dei fatti sia l'essenza della scienza economica che non dev'essere semplice speculazione logica approdo della *l-verità*, ma spiegazione della realtà empirica scopo e identità della *f-verità* come dice l'emistichio di Vico della *Scienza nuova*: « L'ordine dell'idee dee procedere secondo l'ordine delle cose » cui si informa il Demaria di *Logica della produzione e della occupazione*: o meglio spiegazione del mondo sensibile, non del mondo di carta per dirla col Galilei del *Dialogo sopra i due massimi sistemi del mondo* che conferma quella dicotomia – sistema logico e logica dei sistemi – riproposta da Bagioti nel suo *Corso di economia* su cui – se è consentito un peccato d'orgoglio rimessoci forse dalla tradizione ch'è l'anima di ogni scuola – cerchiamo di crescere la nostra *Critica dei massimi sistemi dell'economia politica*.

